Game Theoretic Approaches to Spectrum Sharing in Decentralized Self-Configuring Networks

Samir Perlaza

TELECOM ParisTech

Doctoral School of Informatics, Telecommunications and Electronics
(EDITE)

T H E S I S

presented in partial fulfillment of the Requirements for the Degree of Doctor of Philosophy from TELECOM ParisTech.

Specialization : COMMUNICATIONS AND ELECTRONICS

Samir M. PERLAZA

Game Theoretic Approaches to Spectrum Sharing in Decentralized Self-Configuring Networks

Members of the Jury:

President :	Prof. Philippe CIBLAT, Telecom ParisTech (Paris, France).
Reviewers:	Prof. Eitan ALTMAN, INRIA (Sophia Antipolis, France).
	Prof. David GESBERT, Institut Eurecom (Sophia Antipolis, France).
Examiners:	Prof. Mérouane DEBBAH , SUPELEC (Gif Sur Yvette, France).
	Prof. Robert HEATH, University of Texas at Austin (Austin, TX, USA).
Guest:	M. Guy SALINGUE, Orange Labs (Paris, France).
Thesis Supervisors:	Prof. Pierre DUHAMEL, CNRS-LSS-SUPELEC (Gif Sur Yvette, France).
	Prof. Samson LASAULCE , CNRS-LSS-SUPELEC (Gif Sur Yvette, France).

Thesis defended the 8th of July 2011, in Paris, France.

TELECOM ParisTech (ENST)

École Doctorale Informatique, Télécommunication et Électronique - (EDITE)

T H È S E

présentée pour obtenir le titre de Docteur de Télécom ParisTech.

Spécialité : Communications et Électronique

Samir M. PERLAZA

Le Partage du Spectre dans les Réseaux Décentralisés Auto-Configurables : Une approche par la Théorie des Jeux.

Membres du Jury:

Président	Prof. Philippe Ciblat, Télécom ParisTech (Paris, France).
Rapporteurs	Prof. Eitan Altman, INRIA (Sophia Antipolis, France). Prof. David Gesbert, Institut Eurecom (Sophia Antipolis, France).
Examinateurs	Prof. Mérouane Debbah , SUPELEC (Gif Sur Yvette, France). Prof. Robert Heath, Université du Texas à Austin (Austin, TX, USA).
Invité	M. Guy Salingue, Orange Labs (Paris, France).
Directeurs de thèse	Prof. Pierre Duhamel, CNRS-LSS-SUPELEC (Gif Sur Yvette, France). Prof. Samson Lasaulce , CNRS-LSS-SUPELEC (Gif Sur Yvette, France).

Thèse soutenue le 8 julliet 2011 à Paris, France.

Abstract

This thesis deals with the analysis and design of techniques for spectrum sharing in decentralized self-configuring networks (DSCN). For such networks, spectrum sharing can be broadly divided into two consecutive phases that radio devices implement autonomously, and often individually. In the first phase, radio devices identify their available spectrum access opportunities (SAO). For instance, unused frequency bands, time slots, or spatial directions over which their transmissions are allowed to take place. In the second phase, radio devices determine the optimal transmit/receive configurations for exploiting the available SAOs while guaranteeing a reliable communication. Here, such configurations are described in terms of power allocation policies, modulation-coding schemes, scheduling policies, decoding orders, etc.

For the first phase, we introduce a novel notion of SAO. The main idea consist in exploiting the unused spatial directions (SD) associated with the singular values of the channel matrix of a given primary link using a water-filling power allocation (PA) policy. The method proposed in this thesis for exploiting such opportunities is called opportunistic interference alignment (OIA) and relies on the existence of multiple antennas at both transmitters and receivers. This novel spectrum sharing technique is particularly useful in highly dense networks where classical SAOs such as unused time slots and/or unused frequency bands are short-lasting rare events.

For the second phase, it is well known that the main problem lies in the mutual interference arising from the simultaneous exploitation of the same set of available SAOs. Therefore, to study such a competitive interaction between the radio devices, we make use of tools from game theory. Within this framework, we adopt a particular network topology (parallel multiple access channel) to study the existence and the multiplicity of Nash equilibria (NE). The relevance of NE stems from the fact that it represents a network state where each radio device's configuration is optimal with respect to the configuration of all the other devices. In particular, we show that, paradoxically, an important gain in the global performance at the NE is observed by reducing the number of possible configurations a radio device is allowed use. Later, we introduce a novel technique that allows radio devices to achieve NE in a fully decentralized fashion based only on the periodical observation of their individual performance. This (learning) technique is independent of both the network topology and the performance metric of the radio devices. More importantly, we show that it converges to epsilon-NE in relevant types of games in wireless communications, namely potential games among others.

We finally tackle the quality of service provisioning in DCSNs. We thus formalize an alternative notion of equilibrium, namely satisfaction equilibrium (SE). Contrary to the existing equilibrium concepts, we show that the SE properly models the problem of QoS in DCSNs. More importantly, we introduce learning techniques that allow achieving a SE relying only on a periodical 1-bit message from the receivers. In particular, as long as the SE exists, these techniques achieve a SE in finite time and are shown to be computationally simpler than those used to achieve NE.

Résumé

Les travaux de cette thèse s'inscrivent tous dans la thématique traitement du signal pour les réseaux de communications distribués. Le réseau est dit distribué au sens de la décision. Dans ce cadre, le problème générique et important que nous avons approfondi est le suivant. Comment un terminal, qui a accès à plusieurs canaux de communications, doit-il répartir (de manière autonome) sa puissance d'émission entre ses canaux et l'adapter dans le temps en fonction de la variabilité des conditions de communications? C'est le problème de l'allocation de ressources adaptative et distribuée. Nous avons développé 4 axes de travail qui ont tous conduits à des réponses originales à ce problème; la forte corrélation entre ces axes est expliquée dans le manuscrit de thèse.

Le premier axe a été l'alignement opportuniste d'interférence. Un des scénarios de référence est le cas où deux couples émetteur-récepteur communiquent en interférant (sur la même bande, en même temps, au même endroit, *etc.*), où les 4 terminaux sont équipés de plusieurs antennes et où un émetteur est contraint de ne pas (ou peu) interférer sur l'autre (canal à interférence dit MIMO). Nous avons conçu une technique d'émission de signal multi-antennaire qui exploite l'observation-clé suivante et jamais exploitée auparavant: même lorsqu'un émetteur est égoïste au sens de ses performances individuelles, celui-ci laisse des ressources spatiales (dans le bon espace de signal et que nous avons identifié) vacantes pour l'autre émetteur. L'apport en performances en termes de débit par rapport aux algorithmes existants a été quantifié grâce à la théorie des matrices aléatoires et des simulations Monte Carlo. Ces résultats sont particulièrement importants pour le scénario de la radio cognitive en milieu dense.

Dans un second temps, nous avons supposé que tous les émetteurs d'un réseau sont libres d'utiliser leurs ressources de manière égoïste. Les ressources sont données ici par les canaux fréquentiels et la métrique individuelle de performance est le débit. Ce problème peut être modélisé par un jeu dont les joueurs sont les émetteurs. Une de nos contributions a été de montrer que ce jeu est un jeu de potentiel, ce qui est fondamental pour la convergence des algorithmes distribués et l'existence d'équilibre de Nash. De plus, nous avons montré l'existence d'un paradoxe de Braess: si l'espace d'optimisation d'un joueur grandit, les performances individuelles et globales peuvent s'en trouver réduites. Cette conclusion a une conséquence pratique immédiate: il peut y a voir intérêt de restreindre le nombre de canaux fréquentiels utilisables dans un réseau à interférence distribué.

Dans le jeu précédent, nous avions constaté que les algorithmes distribués d'allocation de ressources (les algorithmes d'apprentissage par renforcement typiquement) demandent un grand nombre d'itérations pour converger vers un état stable tel qu'un équilibre de Nash. Nous avons ainsi proposé un nouveau concept de solution d'un jeu, à savoir l'équilibre de satisfaction; les joueurs ne modifient pas leur action, même si celle-ci ne maximise pas leur gain, pourvu qu'un niveau minimal de performance soit atteint. Nous avons alors développé une méthodologie d'étude de cette solution (existence, unicité, convergence,). Une de nos contributions a aussi été de donner des

algorithmes d'apprentissage qui convergent vers cette solution en un temps fini (et même court génériquement). De nombreux résultats numériques réalisés dans des scénarios imposés par Orange ont confirmé la pertinence de cette nouvelle approche. Le quatrième axe de travail a été la conception de nouveaux algorithmes d'apprentissage qui convergent vers des solutions de type équilibre logit, epsilon-équilibre ou équilibre de Nash. Notre apport a été de montrer comment modifier les algorithmes existants pour que ceux-ci évitent les phénomènes de cycles et convergent vers un équilibre présélectionné au départ de la dynamique. Une idée importante a été d'introduire une dynamique d'apprentissage de la fonction métrique de performances en couplage avec la dynamique principale qui régit l'évolution de la distribution de probabilité sur les actions possibles d'un joueur. Le cadre de ces travaux est parfaitement réaliste d'un point de vue informatif au niveau des terminaux en pratique. Il est montré une voie possible pour améliorer l'efficacité des points de convergence, ce qui constitue un problème encore ouvert dans ce domaine.

Présentation en Français

Dans ce chapitre, nous présentons, en français, une description générale du problème de partage du spectre dans les réseaux décentralisés auto-configurables (decentralized self-configuring network, DSCN). En particulier, nous identifions les principaux défis technologiques et nous décrivons les principales contributions de cette thèse. Enfin, nous exposons nos conclusions et perspectives.

Contexte et Défis Technologiques

Un système de communications décentralisé auto-configurable est essentiellement un réseau sans infrastructure fixe où les appareils radio sont autonomes et déterminent eux mêmes leurs propres configurations de transmission/réception afin de garantir une communication fiable. Pour atteindre cet objectif, les dispositifs radio sont souvent équipés avec des systèmes de détection de spectre et d'auto-configuration. En conséquence, ce type d'appareils radio, souvent appelés radios cognitives (CR), sont capables d'identifier les ressources radio inutilisées et d'adapter leur configurations de transmission/réception pour exploiter plus efficacement ces ressources. Cette configuration est typiquement décrite en termes des politiques d'allocation de puissance, sélection de canaux, schémas de modulation-codage, etc.

La caractéristique à souligner dans le contexte du DCSN est le fait que les émetteurs communiquent avec leurs récepteurs respectifs, sans le contrôle d'une autorité centrale, par exemple, une station de base. Ainsi, la principale limitation de ces réseaux est l'interférence mutuelle résultante de l'interaction sans aucune coordination d'un groupe d'émetteurs exploitant un ensemble commun de ressources radio. En effet, c'est la raison pour laquelle l'analyse et la conception de techniques de partage du spectre joue un rôle central dans ce scénario. Ici, parmi toutes les contraintes pour parvenir à une exploitation optimale du spectre, on mentionne deux faits. Tout d'abord, les appareils radio doivent déterminer leurs propres configurations de transmission/réception basés sur des informations uniquement locales. Deuxièmement, la topologie du réseau est en constant changement. En plus, ces réseaux doivent être rapidement déployables, ainsi que robustes aux attaques et aux pannes dans des environnements critiques. Ces dernières exigences proviennent du fait que les DSCN sont typiquement utilisés dans de nombreuses applications militaires, de secours et commerciales.

En général, le partage des fréquences dans le contexte des DSCN pourrait suivre deux ensembles de règles différentes selon les bandes de fréquences où ils opèrent,

par exemple, des bandes d'accès libre ou des bandes d'accès restreint [122]. Dans le cas des bandes d'accès restreint, nous disons que le DCSN suit un accès hiérarchique au spectre (hierarchical spectrum access, HSA) et dans le cas des bandes d'accès libre, on dit qu'il suit un accès ouvert au spectre (open spectrum access OSA).
En HSA, les DCSNs opérèrent sous la condition que l'interférence supplémentaire produite sur les systèmes pré-existants peut être considérée comme inexistante ou en dessous d'un seuil spécifique [17]. Un exemple typique de HSA est précisément l'idée derrière la radio cognitive, où seulement les ressources spectrales laissées libres par les systèmes préexistants (opportunités d'accès au spectre) sont utilisées, et donc, aucune interférence supplémentaire n'est générée. Dans la suite, nous appelons ces ressources libres : opportunités d'accès au spectre (spectrum access opportunities, SAO). Selon la technique d'accès multiple du système primaire, une SAO peut se concevoir comme une période de temps dans l'accès multiple par répartition temporelle (time division multiple access, TDMA), une bande de fréquence dans l'accès multiple par répartition en fréquence (frequency division multiple access, FDMA), une direction spatiale dans l'accès multiple par répartition spatiale (space division multiple access, SDMA), un ton de fréquence dans l'accès multiple à répartition en fréquences orthogonales (orthogonal frequency division multiple access, OFDMA), un code d'étalement dans l'accès multiple par répartition en code (code division multiple access, CDMA) ou une quelconque combinaison de ceux-ci. Un autre exemple classique de HSA est la modulation en bande ultra-large (ultra-wide band, UWB). Ici, les interférences produites par les appareils radio en utilisant la modulation UWB ne représentent pas une interférence significative supplémentaire aux anciens systèmes, et donc, une telle coexistence est tolérée.
En OSA, chaque appareil radio a les mêmes droits d'accès au spectre à tout moment. Ceci est particulièrement le cas des bandes sans licence (par exemple, la bande pour les applications industrielles, scientifiques et médicales (ISM) [2,400, 2,500] GHz). Les appareils radio fonctionnant dans ces bandes, sont par exemple, des téléphones sans fil, des capteurs sans fil et les appareils fonctionnant sous les normes du Wi-Fi (IEEE 802.11), Zig-Bee (IEEE 802.15.4), et le Bluetooth (IEEE 802.15.1).
Dans cette thèse, nous nous concentrons sur les deux cas, HSA et OSA. En particulier, nous soulignons le fait que dans le HSA, une fois les SAO disponibles sont identifiées de façon fiable ou le niveau d'interférence instantanée produit sur le système primaire est connu par tous les appareils radio dans le DCSN, l'analyse du partage de spectre est identique dans les deux cas HSA et OSA. La remarque importante ici est que, deux ou plusieurs appareils radio qui exploitent les mêmes ressources radio sont soumis aux interférences mutuelles, indépendamment du fait qu'ils opèrent sous politiques de HSA ou SAO. La différence principale entre HSA et OSA est que dans le premier cas, les ressources radio sont disponibles uniquement pendant la période où le système primaire ne les utilise pas, tandis que dans le deuxième cas, les ressources radio sont toujours accessibles. Cependant, dans cette thèse, nous ne considérons pas cette contrainte, et nous supposons que les ressources radio disponibles, en HSA, restent disponibles une période plus longue que la durée de la communication dans la DSCN. Sous cette hypothèse, le problème commun avec

HSA et OSA peut s'exprimer tout simplement comme un groupe d'appareils radio qui exploitent simultanément les mêmes ressources radio et donc, ils sont soumis à des interférences mutuelles. Dans la suite, nous utilisons le terme générique de partage de spectre pour désigner les deux scénarios HSA et OSA dans les conditions indiquées ci-dessus.

Dans ce contexte, le problème d'accès au spectre, comme traité dans cette thèse, peut être décrit par l'ensemble des questions suivantes: (*i*) Quel est la performance optimale individuelle des appareils radio qui peut être observée dans un DSCN? (*ii*) Quel est le comportement optimal qu'un dispositif radio doit adopter pour atteindre une performance optimale individuelle étant donné l'environnement et la configuration de tous les autres appareils?

Pour faire face à la première question, la théorie des jeux, une branche des mathématiques qui étudie les interactions entre plusieurs preneurs de décisions, est le paradigme dominant suivi dans cette thèse [46, 50, 51, 79, 106]. En particulier, nous utilisons l'idée d'équilibre [46] pour déterminer fondamentalement les états stables d'un DSCN donné. Ici, la stabilité est interprétée comme un état où la configuration de transmission/réception de chaque appareil radio est optimale par rapport aux configurations de l'ensemble des autres dispositifs radio. Ainsi, aucun des dispositifs radio améliore ses performances en déviant de façon unilatérale de l'état d'équilibre. Concernant la question (*ii*), par le terme comportement, nous nous référons à la politique qu'un appareil radio utilise pour sélectionner la configuration de transmission/réception en fonction de l'information disponible. Ici, une remarque importante est que certains équilibres peuvent être obtenus comme le résultat d'un processus itératif d'interaction, similaire à un processus d'apprentissage. Ainsi, comme un outil supplémentaire pour faire face à la question (*ii*), nous utilisons certains éléments de la théorie de l'apprentissage de machines.

Les contributions faites dans cette thèse peuvent être classifiées dans trois domaines principaux: (*i*) analyse des performances et conception de techniques d'accès hiérarchique au spectre (HSA), (*ii*) analyse des performances et conception de techniques d'accès ouvert au spectre (OSA), et (*iii*) mécanismes pour l'approvisionnement de la qualité de service en HSA et OSA.

Dans le contexte du HSA, deux contributions sont présentées. La première concerne un système opportuniste de l'alignement d'inférence dans les réseaux MIMO cognitives [77, 78, 87]. La deuxième contribution concerne une technique pour améliorer l'efficacité spectrale des réseaux cognitifs en utilisant une modification stratégique du nombre de canaux que les appareils radio sont autorisés à utiliser [75].

Dans le contexte du OSA, deux contributions sont présentées dans cette thèse. Premièrement, l'analyse des équilibres (de Nash) d'un canal à accès multiple décentralisé et en parallèle [73,80,86]. Ce scénario suppose le cas où plusieurs émetteurs sont destinés à communiquer avec le récepteur à la plus haute efficacité spectrale possible au même temps qu'ils partagent un ensemble commun de bandes de fréquences disponibles. La deuxième contribution consiste en une méthodologie pour la conception de règles de comportement qui permettent aux appareils radio d'atteindre un équilibre comme le résultat d'une interaction itérative similaire à un processus

d'apprentissage [81, 82, 84].
Dans le contexte de l'approvisionnement de qualité de service, la contribution principale consiste en la formalisation d'un concept d'équilibre particulier, à savoir l'équilibre de satisfaction (SE). Contrairement aux notions d'équilibre existants, par exemple l'équilibre de Nash (Nash equilibrium, NE) ou le NE généralisé (generalized Nash equilibrium, GNE), dans la SE, l'idée d'optimisation des performances dans le sens de la maximisation d'une utilité ou la minimisation d'un coût n'existe pas. Le concept de SE repose sur le fait que les joueurs peuvent être satisfaits ou insatisfaits de leur performance. Au SE, si il existe, tous les joueurs sont satisfaits. Cette notion d'équilibre est très bien adaptée pour la modélisation du problème de l'approvisionnement décentralisé de qualité de service dans les DSCN [85].

D'autres contributions faites dans cette thèse consistent en quelques applications des résultats théoriques présentés dans les chapitres à venir. Ces contributions ont été publiées en collaboration avec d'autres auteurs dans le cadre de collaborations avec d'autres laboratoires. Le reste des contributions faites dans cette thèse ont été conservées sous la forme de brevets et elles sont propriété de France Télécom.
Dans la suite, nous décrivons, avec un plus haut niveau de détail, les contributions plus importantes de cette thèse.

Alignement Opportuniste d'Interférence

Un lien multi-entrées-multi-sorties (MIMO) sans interférences avec une connaissance parfaite de l'état du canal à l'émetteur et au récepteur peut être rendu équivalent à plusieurs sous-canaux orthogonaux où ses gains sont les valeurs propres de la matrice de transfert du canal [108]. En utilisant ce modèle équivalent, il est possible d'atteindre la capacité de Shannon en mettant en oeuvre l'allocation de puissance (AP) nommée water-filling [23] entre les différents sous-canaux équivalents. Cependant, les limitations de puissance mènent généralement les émetteurs primaires à laisser certains de ses sous-canaux inutilisés. En fait, les sous-canaux inutilisés, appelés dorénavant ressources spatiales, peuvent donc être réutilisées ou recyclées par un autre système fonctionnant sur la même bande de fréquences [76, 78]. Pour profiter de ces ressources, un certain schéma de construction de signal est exigé : l'émetteur secondaire doit "aligner" son interférence avec les sous-canaux inutilisés de l'émetteur primaire. Il faut noter que dans le domaine fréquentiel, cet alignement s'atteint très simplement avec la transformée de Fourier qui représente une base de décomposition fréquentielle universelle. Avec cela les utilisateurs opportunistes peuvent tout simplement identifier les différentes bandes de fréquences et transmettre à travers celles qui se trouvent libres. Au contraire, dans le domaine spatial, il n'existe pas de base de décomposition spatiale universelle pour tous les utilisateurs. Donc, les utilisateurs secondaires sont sensés connaître le canal du système primaire (au minimum) et traiter leurs signaux pour les aligner avec les mêmes directions spatiales de ce dernier. Les premiers pas vers le concept d'alignement d'interférence sont décris dans [20, 32, 53, 116].

Dans cette thèse, nous proposons une nouvelle technique de construction de signal pour recycler les ressources spatiales. C'est-à-dire, une nouvelle technique d'alignement d'interférence qui exploite les directions spatiales inutilisées par une liaison primaire qui vise à maximiser son débit. Nous proposons également un schéma d'allocation de puissance qui maximise le débit opportuniste. Cette technique est appelée alignement d'interférence car chaque ressource spatiale du système primaire peut être interprétée comme une direction de l'espace. Le lien secondaire doit donc aligner son interférence avec les directions inutilisées. Cette technique est aussi appelée opportuniste car elle profite des limitations de puissance du système primaire et de la réalisation du canal qui l'obligent à concentrer sa puissance dans quelques directions et laisser quelques unes libres.

Cette section est organisée comme suit. Dans la première partie, la conception du système primaire qui vise a maximiser sont débit est décrite. La deuxième partie traite le système opportuniste, plus précisément, le traitement du signal requis au émetteur pour aligner son interférence avec les directions inutilisées du système primaire. Nous décrivons également, l'allocation optimale de puissance. La troisième partie se concentre sur l'estimation du débit asymptotique du système opportuniste. Les conclusions de cette étude sont présentées dans la dernière section.

Modélisation du Système

Nous considérons deux liaisons MIMO point-à-point unidirectionnelles fonctionnant simultanément sur la même bande de fréquence et donc sujettes à des interférences mutuelles. Les liaisons sont supposées indépendantes et non-coopératives, c'est-à-dire qu'aucun échange de messages entre les deux émetteurs n'a lieu avant ou pendant la transmission. Chaque émetteur envoie des messages privés à son récepteur respectif uniquement. Dans notre modèle, les deux émetteurs et les deux récepteurs sont respectivement équipés de N_t antennes et N_r antennes. La première paire émetteur-récepteur, Tx_1 et Rx_1, est la liaison primaire autorisée à exploiter une bande de fréquence donnée de manière exclusive. La paire $\mathrm{Tx}_2 - \mathrm{Rx}_2$ est une liaison opportuniste pouvant exploiter la même bande de fréquence à la condition stricte qu'aucune interférence ne doit être produite sur la liaison primaire. Chaque émetteur est limité en puissance moyenne par un niveau maximal noté $p_{i,\max}$ pour l'émetteur i. Dans cette étude, nous considérons que les deux émetteurs sont limités par le même niveau de puissance $p_{\max}$, c'est-à-dire $\forall i \in \{1,2\}$, $p_{i,\max} = p_{\max}$.

La matrice de transfert du canal entre l'émetteur $j \in \{1,2\}$ et le récepteur $i \in \{1,2\}$ est une matrice $N_r \times N_t$, notée $\mathbf{H}_{ij}$, dont les éléments sont des variables aléatoires complexes et circulaires indépendantes et identiquement distribuées (i.i.d.) selon une loi Gaussienne de moyenne nulle et de variance $\frac{1}{N_t}$. Les matrices de transfert des canaux sont supposées statiques pendant toute la durée de la transmission. Le vecteur regroupant les ζ_i symboles transmis par l'émetteur i est noté $\mathbf{s}_i = (s_{i,1}, \dots, s_{i,\zeta_i})$. Dans notre modèle, l'émetteur i précode linéairement ses symboles en utilisant une matrice $N_t \times \zeta_i$ notée $\mathbf{V}_i$. Dans le cas de la liaison primaire, $\mathbf{V}_1$ est utilisé pour maximiser son débit. Pour la liaison secondaire, $\mathbf{V}_2$ est utilisé pour effectuer l'alignement d'interférence. La variable ζ_i, avec $i \in \{1,2\}$ est décrite dans

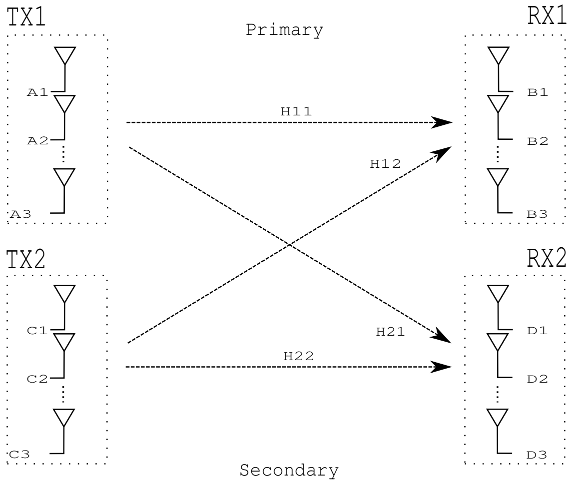

Figure 1: Canal à interférence avec entrées et sorties multi dimensionelles (MIMO).

la section Sec. . Les signaux $\mathbf{r}_1$ et $\mathbf{r}_2$ reçus par les récepteurs primaire et secondaire s'écrivent respectivement

$$\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{V}_1\mathbf{s}_1 \\ \mathbf{V}_2\mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix}, \tag{1}$$

où $\mathbf{n}_i$ est un vecteur de dimension N_r représentant les effets du bruit thermique au récepteur i, dont les éléments sont modélisés par un processus aléatoire Gaussien complexe de moyenne nulle et de matrice de covariance $\mathbb{E}\left[\mathbf{n}_i\mathbf{n}_i^H\right] = \sigma_i^2\mathbf{I}_{N_r}$, $\forall i \in \{1,2\}$. La matrice d'allocation de puissance $\mathbf{P}_i$, de taille $\zeta_i \times \zeta_i$, est définie comme la matrice de covariance $\mathbf{P}_i = \mathbb{E}\left[\mathbf{s}_i\mathbf{s}_i^H\right]$. Nous supposons les contraintes de puissance suivantes:

$$\forall i \in \{1,2\}, \quad \text{Trace}\left(\mathbf{V}_i\mathbf{P}_i\mathbf{V}_i^H\right) \leqslant N_t p_{\max}. \tag{2}$$

À chaque récepteur i, les signaux reçus $\mathbf{r}_i$ sont traités par une matrice de taille $N_r \times N_r$, notée $\mathbf{D}_i$. Ainsi, le signal au récepteur après traitement, noté $\mathbf{y}_i$, est représenté par un vecteur de dimension N_r défini comme $\mathbf{y}_i = \mathbf{D}_i\mathbf{r}_i$ $\forall i \in \{1,2\}$.

Nous décrivons la configuration de la liaison primaires dans la section suivante. La configuration de la liaison opportuniste est décrite dans la section Sec. .

Conception de la Liaison Primaire

Le système primaire est modélisé par une liaison MIMO $N_t \times N_r$ sans interférences. La stratégie optimale d'allocation de puissance pour ce modèle a été étudiée par Telatar [108]. Nous décrivons une telle stratégie par le théorème suivant.

Théorème 0.0.1 (Telatar-1995 [108]) Soit $\mathbf{H}_{11} = \mathbf{U}_{H_{11}}\mathbf{\Lambda}_{H_{11}}\mathbf{V}_{H_{11}}^H$ avec

$$\mathbf{\Lambda} = \text{diag}\left(\lambda_{H_{11},1}, \ldots, \lambda_{H_{11},\zeta_2}\right),$$

la décomçposition en valeurs singulières de la matrice de transfert du canal $\mathbf{H}_{11}$ de dimensions $N_r \times N_t$. La liaison primaire atteint la capacité de Shannon en utilisant la configuration $\mathbf{V}_1 = \mathbf{V}_{H_{11}}$, $\mathbf{D}_1 = \mathbf{U}_{H_{11}}^H$, $\mathbf{P}_1 = \mathrm{diag}(p_{1,1}, \ldots, p_{1,N_t})$, où

$$\forall i \in \{1, \ldots, N_t\}, \quad p_{1,i} = \left[\beta - \frac{\sigma_1^2}{\lambda_{H_{11}^H H_{11},i}}\right]^+, \tag{3}$$

avec, $\mathbf{\Lambda}_{H_{11}^H H_{11}} = \mathbf{\Lambda}_{H_{11}}^H \mathbf{\Lambda}_{H_{11}} = \left(\lambda_{H_{11}^H H_{11},1}, \ldots, \lambda_{H_{11}^H H_{11},N_t}\right)$. La constante β est déterminée pour satisfaire la condition (2).

Les puissances de transmission (3) peuvent être déterminées de manière itérative en utilisant l'algorithme d'allocation de puissance nommé water-filling [23].

Conception de la Liaison Secondaire

Dans cette section, le fonctionnement de la liaison secondaire est décrit. Auparavant, on assume que les valeurs propres de toutes les matrices sont notées par ordre decroissant. C'est-à-dire, si une matrice quelconque, notée $\mathbf{X}$ a N valeurs propres, nous les notons $\lambda_{X,1}, \ldots, \lambda_{X,N}$ et en plus, $\lambda_{X,1} \geqslant \lambda_{X,2} \geqslant \ldots \geqslant \lambda_{X,N}$.

En accord, avec l'idée initiale de ne pas produire de l'interférence dans le système primaire, nous considérons que le système secondaire doit fonctionner sous la contrainte suivante:

Définition 0.0.1 (Condition d'Alignement d'Interférence) Le système secondaire satisfait la condition d'alignement d'interférence (AI) si le système primaire atteint le débit qu'il atteindrait quand le système secondaire ne transmet pas. Nous exprimons cette condition plus formellement comme

$$\begin{aligned} \log_2 \left|\mathbf{I}_{N_r}\sigma_1^2 + \mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H\right| - \log_2 \left|\mathbf{I}_{N_r}\sigma_1^2\right| = \\ \log_2 \left|\mathbf{R} + \mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H\right| - \log_2 \left|\mathbf{R}\right| \end{aligned} \tag{4}$$

où la matrice $\mathbf{R} \triangleq \sigma_1^2 \mathbf{I}_{N_r} + \mathbf{U}_{H_{11}}^H \mathbf{H_{12}} \mathbf{V}_2 \mathbf{P}_2 \mathbf{V}_2^H \mathbf{H}_{12}^H \mathbf{U}_{H_{11}}$ est la matrice de covariance du signal d'interference produite par l'émetteur secondaire ajoutée au bruit du recepteur primaire.

La condition suffisante d'alignement d'interférence est satisfaite si la matrice de pré-traitement $\mathbf{V}_2$ est telle que

$$\mathbf{H}_{12}\mathbf{V}_2 = \mathbf{0}_{N_r \times \zeta_2}. \tag{5}$$

Cette solution est connue comme formation de faisceaux à forçage de zéro (Zero-Forcing beamforming) [72]. Cependant, cette solution n'exploite pas le fait que la liaison primaire laisse inutilisées certaines directions spatiales à cause de ses limitations de puissance de transmission. En fait, chaque direction inutilisée du système primaire peut être interprétée comme une opportunité additionnelle de transmission pour le système secondaire.

Définition 0.0.2 (Opportunités de Transmission (OT) supplémentaires) *Nous disons que le système opportuniste a S opportunités de transmission (OT) s'il existe un ensemble $\mathcal{S} \subset \{1, \ldots, \min(N_t, N_r)\}$ tel que $|\mathcal{S}| = S$ et pour tout $s \in \mathcal{S}$, $\lambda_{H_{11}^H H_{11},s} \neq 0$ et $p_{1,s} = 0$.*

Schéma Optimal de Pré et Post-Traitement du Signal

Pour profiter des OT identifiées dans la section précédente, l'émetteur opportuniste doit déterminer sa matrice de pré-traitement $\mathbf{V}_2$ pour satisfaire la condition (Def. 0.0.1) indépendamment de la matrice d'allocation de puissance $\mathbf{P}_2$. Ce résultat est fourni par le théorème suivant:

Théorème 0.0.2 (Matrice optimale de pré-traitement $\mathbf{V}_2$) *Nous considérons la matrice $\tilde{\mathbf{H}} = \mathbf{U}_{H_{11}}^H \mathbf{H}_{12}$ et sa structure de blocs*

$$\tilde{\mathbf{H}} = \begin{array}{r} N_r - S \updownarrow \\ S \updownarrow \end{array} \begin{pmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \end{pmatrix}. \tag{6}$$

La condition d'AI (Def. 0.0.1) est satisfaite indépendamment de la matrice d'allocation de puissance $\mathbf{P}_2$ quand la matrice de pré-traitement $\mathbf{V}_2$ satisfait la condition:

$$\tilde{\mathbf{H}}_1 \mathbf{V}_2 = \mathbf{0}_{N_r \times \zeta_2}. \tag{7}$$

Il est important de remarquer que n'importe quelle solution différente de celle du théorème (Th. 0.0.2) implique une matrice d'allocation de puissance particulière. Dans notre cas, le but est précisément de satisfaire la condition d'AI en ajustant seulement la matrice de pré-traitement. De cette manière, la matrice d'allocation de puissance reste libre pour être ajustée en cherchant la maximisation du débit du système secondaire. La matrice de post-traitement du signal est choisie tout simplement comme un filtre blanchisseur du signal d'entrée. Ce choix est optimal dans le sens qu'il n'existe aucune perte d'information mutuelle entre le signal d'entrée et le signal après le filtrage. Donc,

$$\mathbf{D}_2 = \mathbf{Q}^{-\frac{1}{2}}, \tag{8}$$

où $\mathbf{Q} = \mathbf{H}_{21}\mathbf{V}_{H_{11}}\mathbf{P}_1\mathbf{V}_{H_{11}}^H\mathbf{H}_{21}^H + \sigma_2^2\mathbf{I}_{N_r}$ est la matrice de covariance du signal d'interférence produit par le système primaire ajoutée au bruit du récepteur secondaire.

La section suivante s'occupe du problème d'optimisation qui vise la maximisation du débit opportuniste.

Schéma d'Allocation de Puissance

Le problème d'intérêt dans cette section peut être écrit comme:

$$\begin{array}{ll} \max\limits_{\mathbf{P}_2} & \log_2 \left| \mathbf{I}_{N_r} + \mathbf{Q}^{-1}\mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H \right| \\ \text{s.t.} & \mathrm{Trace}\left(\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\right) \leqslant p_{\max}. \end{array} \tag{9}$$

Avant de résoudre le problème d'optimisation dans (9), nous décrivons brièvement l'allocation uniforme de puissance (AUP). Dans quelques situations, AUP peut être préférée à la solution optimale (Allocation Optimale de Puissance, AOP) pour sa simplicité de calcul. En effet, pour un petit nombre de OTs, par exemple $S < 3$, le gain en débit obtenu avec AOP n'est pas très significatif par rapport à celui d'AUP.

Allocation Uniforme de Puissance

Dans le cas d'AUP, l'émetteur divise la totalité de sa puissance entre toutes les TOs ayant été identifiées, i.e., $\mathbf{P}_{2,UPA} = \gamma \mathbf{I}_{\zeta_2}$ où

$$\gamma = \frac{N_t\, p_{\max}}{\mathrm{Trace}\left(\mathbf{V}_2 \mathbf{V}_2^H\right)}. \tag{10}$$

Allocation Optimale de Puissance

La puissance de transmission qui maximise le débit du système secondaire, i.e. la solution au problème d'optimisation (9), est aussi une allocation de puissance sous la forme du water-filling.

Théorème 0.0.3 (Allocation Optimal de Puissance) *Nous considérons la matrice* $\mathbf{K} \triangleq \mathbf{Q}^{-\frac{1}{2}} \mathbf{H}_{22} \mathbf{V}_2$ *et sa décomposition en valeurs singulières* $\mathbf{K} = \mathbf{U}_K \mathbf{\Lambda}_K \mathbf{V}_K^H$, *avec* $\mathbf{\Lambda}_K = \mathrm{diag}\left(\lambda_{K,1}, \ldots, \lambda_{K,\zeta_2}\right)$. *La matrice d'allocation optimale de puissance est*

$$\mathbf{P}_2 \;=\; \mathbf{V}_K \tilde{\mathbf{P}} \mathbf{V}_K^H, \tag{11}$$

où $\tilde{\mathbf{P}} = \mathrm{diag}\left(\tilde{p}_1, \ldots, \tilde{p}_{\zeta_2}\right)$ *est une matrice diagonale avec entrées données par*

$$\forall i \in \{1, \ldots, \zeta_2\}, \quad \tilde{p}_{2,i} = \left[\beta_o - \frac{1}{\lambda_{K^H K,i}}\right]^+. \tag{12}$$

La matrice $\lambda_{\mathbf{K^H K}} = \lambda_{\mathbf{K^H}} \lambda_{\mathbf{K}} = \mathrm{diag}\left(\lambda_{K^H K,1}, \ldots, \lambda_{K^H K,\zeta_2}\right)$ et β_o est une constante qui satisfait les contraintes de puissance du système secondaire (2).

La Fig. 2 montre les débits atteignables du système secondaire pour un nombre arbitraire d'antennes quand $N_t = N_r + 1$. Il faut noter que la performance de la technique d'IA est toujour superiéure ou égale à celle de la technique de formation de faisceaux à forçage de zero (Zero-Forcing Beamforming) [72].

Débit Asymptotique de la Liaison Secondaire

Le débit asymptotique du système secondaire peut-être déterminée sous l'hypothèse d'un grand nombre d'antennes, i.e. $N_t, N_r \to \infty$, avec $\frac{N_r}{N_t} = \alpha < \infty$ en utilisant le théorème suivant:

Théorème 0.0.4 (Débit Asymptotique de la Liaison Secondaire) Nous considérons un système primaire et secondaire qui utilisent leurs configurations optimales. Nous assumons que N_r, $N_t \to \infty$, avec $\frac{N_r}{N_t} \to \alpha < \infty$, et $\mathbf{M}_1 \triangleq \mathbf{H}_{12} \mathbf{V}_{H_{11}} \mathbf{P}_1 \mathbf{V}_{H_{11}}^H \mathbf{H}_{12}^H$,

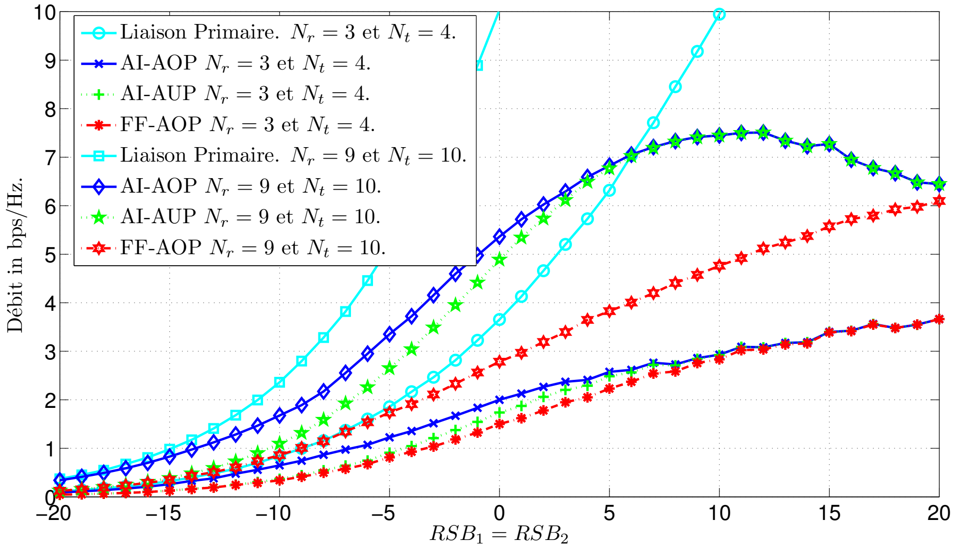

Figure 2: Débit du système opportuniste comme fonction du rapport signal sur bruit $\mathrm{RSB}_1 = \frac{p_{\max}}{\sigma_1^2}$. Le nombre d'antennes satisfait $N_t = N_r + 1$, avec $N_r \in \{3, 9\}$ et $RSB_1 = RSB_2$. La technique de formation de faisceaux (FF) à forçage de zero suit l'equation 5 avec puissance optimale.

$\mathbf{M}_2 \triangleq \mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H$, $\mathbf{M} \triangleq \mathbf{M}_1 + \mathbf{M}_2$. Alors, le débit asymptotique par antenne du système opportuniste (Tx_2-Rx_2) est donnée par

$$\bar{R}_2(p_{\max}, \sigma_2^2) = \frac{1}{\ln 2} \int_{\sigma_2^2}^{+\infty} \mathbf{G}_{M_1}\left(-\sigma_2^2\right) - \mathbf{G}_{M}\left(-\sigma_2^2\right) \mathrm{d}\sigma_2^2, \tag{13}$$

où, $\mathbf{G}_M(z)$ et $\mathbf{G}_{M_1}(z)$ sont les transformées de Stieltjes des distributions empiriques des valeurs propres des matrices $\mathbf{M}$ et $\mathbf{M}_1$, respectivement. Le deux $\mathbf{G}_M(z)$ et $\mathbf{G}_{M_1}(z)$ sont obtenus comme solutions des équations de point fixe (avec solution unique quand $z \in \mathbb{R}_-$), $G_M(z) = \frac{-1}{z - g(G_M(z)) - h(G_M(z))}$, et $G_{M_1}(z) = \frac{-1}{z - g(G_{M_1}(z))}$, respectivement. Les fonctions $g(z)$ et $h(z)$ sont définies comme

$$g(u) \triangleq \mathbb{E}\left[\frac{p_1}{1 + \frac{1}{\alpha} p_1 u}\right], \text{ and} \tag{14}$$

$$h(u) \triangleq \mathbb{E}\left[\frac{p_2}{1 + \frac{1}{\alpha} p_2 u}\right]. \tag{15}$$

$$\tag{16}$$

Dans les expressions 14) et 15 l'espérance est calculée avec la distribution de probabilité des variables p_1 et p_2, i.e., $F_{P_j}(\lambda)$, où

$$\forall j \in \{1, 2\}, \quad F_{P_j}(\lambda) \triangleq \frac{1}{\zeta_j} \sum_{i=1}^{\zeta_j} \mu(\lambda - p_{j,i}). \tag{17}$$

La Fig. 3 montre le débit asymptotique obtenue avec le théorème (Th. 0.0.4) et le débit obtenu en utilisant un grand nombre d'antennes quand $N_r = N_t$. Ces résultats montrent que dans le régime asymptotique, le système opportuniste arrive à atteindre des débits du même ordre de ceux du système primaire. Plus important, la figure montre comme le rapport signal sur bruit (RSB) du système primaire joue un role important dans le débit atteignable du système secondaire.

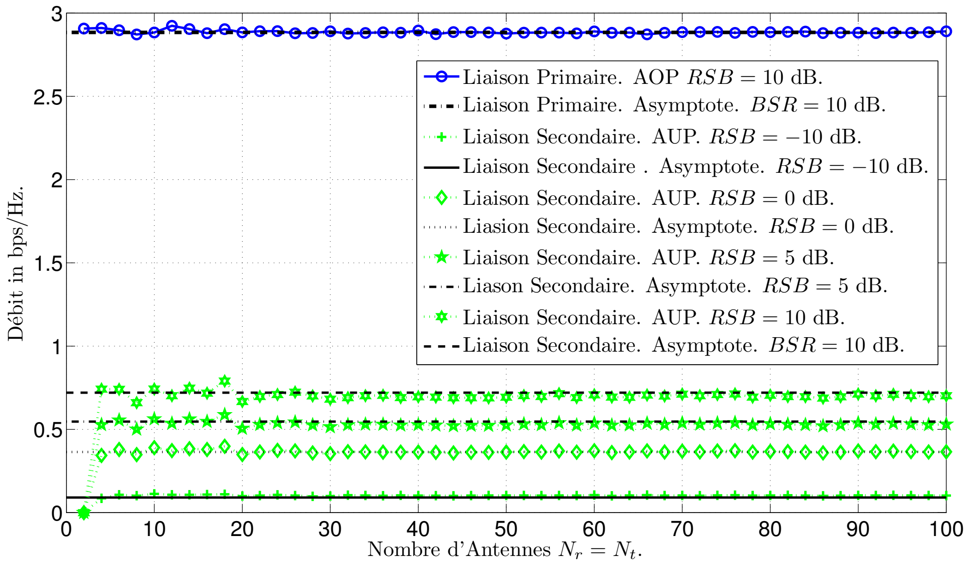

Figure 3: Débit Asymptotique du système opportuniste avec AUP observé par simulation comme fonction du nombre d'antennes quand $N_r = N_t$ à différents niveaux de RSB. RSB $= \frac{p_{\max}}{\sigma^2}$. Les lignes noires sont des asymptotes déterminées par le Th. 0.0.4. Les asymptotes du système primaire sont données par [30].

Conclusions

Nous avons proposé une nouvelle technique qui permet aux liaisons point-à-point MIMO opportunistes de recycler les ressources spatiales laissées inutilisées par des liaisons point-à-point MIMO primaires. Nous avons fourni la technique de construction de signal pour exploiter ces ressources spatiales et aussi un schéma d'allocation de puissance qui maximise le débit des liaisons opportunistes. Une analyse sous l'hypothèse d'un grand nombre d'antennes permet de déterminer asymptotiquement le débit maximal du système secondaire. D'après cette analyse on trouve que le système secondaire est capable d'atteindre des débits de transmission du même ordre que la liaison primaire.

Nash Equilibria in Parallel Multiple Access Channel

Un canal à accès multiple (Multiple Access Channel, MAC) correspond à un scénario dans lequel il y a plusieurs émetteurs et un seul récepteur. Dans le cas du canal à accès multiple en parallèle, chaque émetteur peut exploiter plusieurs canaux orthogonaux pour communiquer avec le récepteur. Ce modèle de canal permet l'analyse de la macro-diversité dans la liaison montante des réseaux cellulaires (dans ce cas, les stations de base sont supposées être connectées à une entité centrale commune), l'allocation de puissance dans les canaux d'accès multiples sélectives en fréquence lorsque un multiplexage par division en fréquences orthogonales (Orthogonal Frequency Division Multiplexing, OFDM) est utilisée, ou la sélection de point d'accès dans les réseaux locaux sans fil. En termes de canaux multi-utilisateurs, le MAC en parallèle correspond à un cas particulier du MAC vectoriel [121], mais ici, le système est supposé décentralisé, c'est-à-dire, tous les émetteurs peuvent choisir librement leur politique d'accès au spectre. Ce choix peut être soit une des politiques d'allocation de puissance (Power Allocation, PA) entre les canaux disponibles ou une politique de sélection de canal (Channel Selection, CS). La métrique de performance pour chaque terminal dans cet étude est son efficacité spectrale individuelle. Nous nous référerons à ces problèmes comme le problème/jeu (a) et problème/jeu (b), respectivement. Les problèmes (a) et (b) peuvent être modélisés par des jeux sous forme stratégique où les joueurs sont les émetteurs, la fonction de paiement/récompense/utilité des joueurs est l'efficacité spectrale individuelle, et l'ensemble des actions est l'ensemble des politiques d'allocation de puissance ou de sélection de canal. Le concept de solution utilisé dans cette thèse est celle de l'équilibre de Nash [66]. Cet équilibre peut être atteint comme le résultat d'un processus dynamique d'apprentissage impliquant des hypothèses d'informations raisonnables, ce qui fait de cette notion la plus adaptée pour le sujet traité dans cette thèse est le fait que. En particulier, cette propriété permet la conception des algorithmes distribués pour l'exploitation du spectre.

Modèle du Jeu

Les problèmes de PA et CS décrits ci-dessus peuvent être modélisés respectivement par les deux jeux sous forme stratégique (avec $i \in \{a, b\}$):

$$\mathcal{G}_{(i)} = \left(\mathcal{K}, \left(\mathcal{P}_k^{(i)}\right)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right). \tag{18}$$

Dans les deux jeux, l'ensemble des émetteurs $\mathcal{K}$ est l'ensemble des joueurs. Une action d'un émetteur donné $k \in \mathcal{K}$ est un schéma particulier de PA, c'est-à-dire un vecteur de PA de dimension S, noté par $\mathbf{p}_k = (p_{k,1}, \ldots, p_{k,S}) \in \mathcal{P}_K^{(i)}$, où $\mathcal{P}_K^{(i)}$ est l'ensemble de tous les vecteurs de PA possibles que l'émetteur k peut utiliser, soit dans le jeux $\mathcal{G}_{(a)}$ $(i = a)$ ou dans le jeu $\mathcal{G}_{(b)}$ $(i = b)$. Nous appelons un vecteur de la forme

$$\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_K) \in \mathcal{P}^{(i)},$$

un profil d'action du jeu $i \in \{a,b\}$ où $\mathcal{P}^{(i)}$ est un ensemble obtenu du produit Cartésien des ensembles des actions $\mathcal{P}_k^{(i)}$, pour tout $k \in \mathcal{K}$, i.e., $\mathcal{P}^{(i)} = \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_K^{(i)}$, où,

$$\begin{aligned} \mathcal{P}_k^{(a)} &= \Big\{(p_{k,1}, \ldots p_{k,S}) \in \mathbb{R}^S : \forall s \in \mathcal{S}, p_{k,s} \geqslant 0, \\ & \quad \sum_{s \in \mathcal{S}} p_{k,s} \leqslant p_{k,\max}\Big\}, \text{ et} \end{aligned} \tag{19}$$

$$\begin{aligned} \mathcal{P}_k^{(b)} &= \{p_{k,\max}\, \mathbf{e}_s : \forall s \in \mathcal{S},\ \mathbf{e}_s = (e_{s,1}, \ldots, e_{s,S}), \\ & \quad \forall r \in \mathcal{S} \setminus s, e_{s,r} = 0, \text{ and } e_{s,s} = 1\}. \end{aligned} \tag{20}$$

Dans la suite, nous nous référons aux jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$ comme le jeu PA et le jeu SC. Nous noterons par $\mathbf{p}_{-k}$ tout vecteur dans l'ensemble

$$\mathcal{P}_{-k}^{(i)} \triangleq \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_{k-1}^{(i)} \times \mathcal{P}_{k+1}^{(i)} \times \ldots \times \mathcal{P}_K^{(i)} \tag{21}$$

avec $(i,k) \in \{a,b\} \times \mathcal{K}$. Pour un $k \in \mathcal{K}$ donné, le vecteur noté par $\mathbf{p}_{-k}$ représente les stratégies adoptés par tous les autres joueurs différents au joueur k. Avec un léger abus de notation, nous écrivons parfois tout vecteur $\mathbf{p} \in \mathcal{P}^{(i)}$ comme $(\mathbf{p}_k, \mathbf{p}_{-k})$, afin de souligner la k-ème composante du super vecteur $\mathbf{p}$. La fonction d'utilité du joueur k dans le jeu $\mathcal{G}_{(i)}$ est l'efficacité spectrale $u_k :\ \mathcal{P}^{(i)} \to \mathbb{R}_+$, et

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) = \sum_{s \in \mathcal{S}} \frac{B_s}{B} \log_2(1 + \gamma_{k,s}) \ \text{[bps/Hz]} \tag{22}$$

où $\gamma_{k,s}$ est le rapport signal sur bruit plus interférence (RSBI) vu par le joueur k dans le canal s, i.e.,

$$\gamma_{k,s} = \frac{p_{k,s} g_{k,s}}{\sigma_s^2 + \sum\limits_{j \in \mathcal{K} \setminus \{k\}} p_{j,s} g_{j,s}}, \tag{23}$$

et $g_{k,s} \triangleq |h_{k,s}|^2$. Nous assumons un décodage du type "single-user decoding" (SUD) en réception. De toute évidence, nous ne cherchons pas l'optimalité du système, mais tout simplement, un choix de décodage qui puisse être escaladé et équitable avec tous les utilisateurs. Les jeux $\mathcal{G}_{(i)}$, $i \in \{a,b\}$ correspondent à un conflit d'intérêt entre décideurs égoïstes. Ici, l'interaction est due à l'interférence d'accès multiple et les décisions consistent dans le choix des vecteurs de PA. À partir de certaines hypothèses sur l'informations et comportement des émetteurs, une question naturelle est de savoir si ce conflit a des résultats prévisibles. Ainsi, suivant ce raisonnement, nous nous concentrons sur l'équilibre de Nash [66] comme un concept de solution de ce conflit. Un NE pure est définie comme suit.

Définition 0.0.3 (Équilibre de Nash Pur) *Dans les jeux non-coopératifs sous forme stratégique $\mathcal{G}_{(i)}$, avec $i \in \{a,b\}$, un profil d'action $\boldsymbol{p} \in \mathcal{P}^{(i)}$ est un NE pur si, pour tout $k \in \mathcal{K}$ et pour tout $\mathbf{p}_k' \in \mathcal{P}_k^{(i)}$, que*

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) \geqslant u_k(\mathbf{p}_k', \mathbf{p}_{-k}). \tag{24}$$

Lorsque au moins un NE existe dans les jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$, il peut être atteint comme le résultat d'une interaction à long terme entre des joueurs qui suivent une règle particulière de comportement (voir Sec.) avec un échange d'information réduit entre le récepteur et les émetteurs.
Les jeux de potentiel (GP) [64] sont une classe de jeux pour lesquels l'existence d'au moins un équilibre pur de Nash est garantie. En outre, de nombreuses procédures d'apprentissage, telles que la dynamique de meilleure réponse, le jeu fictif et certaines dynamiques d'apprentissage par renforcement convergent vers un NE dans les GP. Un des buts de cette thèse est de montrer que les jeux $\mathcal{G}_{(i)}$, $i \in \{a, b\}$ sont des jeux de potentiel [64, 96].

Définition 0.0.4 (Jeux Exact de Potentiel) *Tout jeu sous forme stratégique définie par le triplet* $\left(\mathcal{K}, (\mathcal{P}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right)$ *est un jeu exact de potentiel s'il existe une fonction* $\phi(\mathbf{p})$ *pour tout* $\mathbf{p} \in \mathcal{P} = \mathcal{P}_1 \times \ldots \times \mathcal{P}_K$ *telle que pour tous les joueurs* $k \in \mathcal{K}$ *et pour tous les vecteurs d'allocation de puissance* $\mathbf{p}'_k \in \mathcal{P}_k$*,*

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) - u_k(\mathbf{p}'_k, \mathbf{p}_{-k}) = \phi(\mathbf{p}_k, \mathbf{p}_{-k}) - \phi(\mathbf{p}'_k, \mathbf{p}_{-k}).$$

De la définition de la fonction d'utilité (22), la proposition suivante peut être facilement démontrée [73].

Proposition 0.0.1 *Les jeux sous forme stratégique* $\mathcal{G}_{(i)}$*, avec* $i \in \{a, b\}$*, sont des jeux exacts de potentiel avec fonction de potentiel,*

$$\phi(\mathbf{p}) = \sum_{s\in\mathcal{S}} \frac{B_s}{B} \log_2 \left(\sigma_s^2 + \sum_{k=1}^{K} p_{k,s} g_{k,s}\right). \tag{25}$$

Dans la suite, nous présentons les principaux résultats de cette thèse par rapport au canal d'accès multiple en parallèle en utilisant le cas plus simple. C'est-à-dire, nous ne considérons que deux joueurs et deux canaux pour des raisons de simplicité dans la présentation. Dans les chapitres à venir, des scénarios plus généraux sont étudiés.

Le Cas de 2-Transmetteurs et 2-Canaux

Considérons les jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$ avec $K = 2$ et $S = 2$. Considérons aussi $\forall k \in \mathcal{K}$, $p_{k,\max} = p_{\max}$ et $\forall s \in \mathcal{S}$, $\sigma_s^2 = \sigma^2$ et $B_s = \frac{B}{S}$. Notons par $\mathrm{SNR} = \frac{p_{\max}}{\sigma^2}$ le RSB de chaque lien de communication.

Le Jeux d'Allocation de Puissance

Notons par $\mathbf{p}^\dagger = \left(\mathbf{p}_1^\dagger, \mathbf{p}_2^\dagger\right)$ l'equilibre du jeu $\mathcal{G}_{(a)}$. Ansi, en suivant la Déf. 0.0.3, nous écrivons l'ensemble d'équations suivant,

$$\forall k \in \mathcal{K}, \quad \mathbf{p}_k^\dagger \in \mathrm{BR}_k\left(\mathbf{p}_{-k}^\dagger\right). \tag{26}$$

Il est important de remarquer que pour tout $k \in \mathcal{K}$ et pour tout $\mathbf{p}_{-k} \in \mathcal{P}^{(a)}$, l'ensemble $\mathrm{BR}_k(\mathbf{p}_{-k})$ est un singleton (Déf. 3.3.3) et alors, (26) représente un système d'équations. En résolvant ce système d'équations (26) pour une réalisation donnée des canaux $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$, nous déterminons l'équilibre du jeu $\mathcal{G}_{(a)}$. Nous présentons cette solution dans la proposition suivante.

Proposition 0.0.2 (Équilibre de Nash dans le jeu $\mathcal{G}_{(a)}$) *Notons par* $\mathbf{p}^\dagger = \left(\mathbf{p}_1^\dagger, \mathbf{p}_2^\dagger\right) \in \mathcal{P}^{(a)}$, *avec* $\mathbf{p}_1^\dagger = \left(p_{11}^\dagger, p_{\max} - p_{11}^\dagger\right)$ *et* $\mathbf{p}_2^\dagger = \left(p_{\max} - p_{22}^\dagger, p_{22}^\dagger\right)$ *un des équilibres du jeu* $\mathcal{G}_{(a)}$. *Alors, avec probabilité un,* $\mathbf{p}^\dagger$ *est l'unique équilibre du jeu et peut être écrit comme:*

- *Équilibre* 1: *lorsque* $\mathbf{g} \in \mathcal{B}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}, \frac{g_{21}}{g_{22}} \leqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}\}$, $p_{11}^\dagger = p_{\max}$ *et* $p_{22}^\dagger = p_{\max}$.
- *Équilibre* 2: *lorsque* $\mathbf{g} \in \mathcal{B}_2 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right), \frac{g_{21}}{g_{22}} \geqslant 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right)\}$, $p_{11}^\dagger = p_{\max}$ *et* $p_{22}^\dagger = 0$.
- *Équilibre* 3: *lorsque* $\mathbf{g} \in \mathcal{B}_3 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})}, \frac{g_{21}}{g_{22}} \leqslant \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})}\}$, $p_{11}^\dagger = 0$ *et* $p_{22}^\dagger = p_{\max}$.
- *Équilibre* 4: *lorsque* $\mathbf{g} \in \mathcal{B}_4 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}, \frac{g_{21}}{g_{22}} \geqslant \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}\}$, $p_{11}^\dagger = 0$ *et* $p_{22}^\dagger = 0$.
- *Équilibre* 5: *lorsque* $\mathbf{g} \in \mathcal{B}_5\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{g_{21}}{g_{22}}, \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}} < \frac{g_{21}}{g_{22}} < 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right)\}$, $p_{11}^\dagger = p_{\max}$ *et* $p_{22}^\dagger = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{22}} + \frac{\sigma^2 + g_{11}p_{\max}}{g_{21}}\right)$.
- *Équilibre* 6: *lorsque* $\mathbf{g} \in \mathcal{B}_6\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{g_{21}}{g_{22}}, \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})} < \frac{g_{11}}{g_{12}} < \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}\}$, $p_{11}^\dagger = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{11}} + \frac{\sigma^2 + p_{\max}g_{22}}{g_{12}}\right)$ *et* $p_{22}^\dagger = p_{\max}$.
- *Équilibre* 7: *lorsque* $\mathbf{g} \in \mathcal{B}_7 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{22}}, \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}} < \frac{g_{11}}{g_{12}} < 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right)\}$, $p_{11}^\dagger = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + p_{\max}g_{21}}{g_{11}} + \frac{\sigma^2}{g_{12}}\right)$ *et* $p_{22}^\dagger = 0$.
- *Équilibre* 8: *lorsque* $\mathbf{g} \in \mathcal{B}_8\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{22}}, \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})} < \frac{g_{21}}{g_{22}} < \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}\}$; $p_{11}^\dagger = 0$ *et* $p_{22}^\dagger = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{22}} + \frac{\sigma^2}{g_{21}}\right)$.

Proof: Voir l'annexe D □

Dans la Fig. 4 nous avons tracé les différents types d'équilibres du jeu $\mathcal{G}_{(a)}$ en fonction des rapports de canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$. L'unicité de l'équilibre n'est pas assurée, sous certaines conditions comme nous le montrons dans la proposition suivante. En fait, dans ces conditions un nombre infini d'équilibres peut être observé, toutefois, ces conditions sont observées avec une probabilité nulle.

Proposition 0.0.3 *Supposons que l'ensemble des canaux* $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$ *vérifie les conditions suivantes*

$$\frac{1}{1+\frac{p_{\max}}{\sigma^2}(g_{12}+g_{22})} < \frac{g_{11}}{g_{12}} = \frac{g_{21}}{g_{22}} < 1 + \frac{p_{\max}}{\sigma^2}(g_{11}+g_{21}), \tag{27}$$

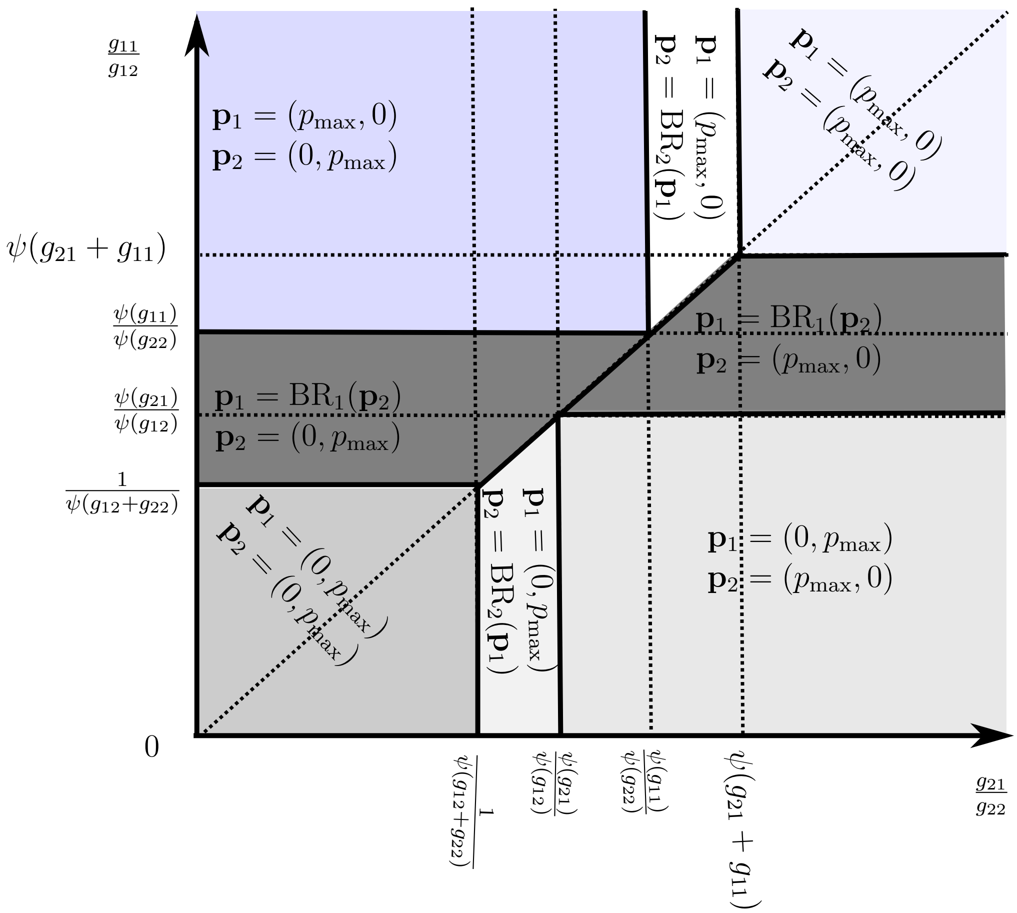

Figure 4: Équilibres de Nash comme fonction des rapports des canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$ pour le cas de deux joueurs et deux canaux dans les jeu $\mathcal{G}_{(a)}$. La fonction $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ est définie comme: $\psi(x) = 1 + \mathrm{SNR}\, x$. La correspondance de meilleure réponse $\mathrm{BR}_k(\mathbf{p}_{-k})$, pour tout $k \in \mathcal{K}$, est définie par (3.14). Dans cet exemple, nous avons choisi arbitrairement $\frac{\psi(g_{21})}{\psi(g_{12})} < \frac{\psi(g_{11})}{\psi(g_{22})}$.

Alors, tout vecteur d'allocation de puissance $\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathcal{P}^{(a)}$*, tel que*

$$p_{11} = \frac{1}{2}\left(p_{\max}(1-\alpha) + \sigma^2\left(\frac{1}{g_{12}} - \frac{1}{g_{11}}\right)\right) + \alpha p_{22}$$

avec $\alpha \triangleq \frac{g_{11}}{g_{21}} = \frac{g_{12}}{g_{22}}$*, est un profil d'équilibre de Nash du jeu* $\mathcal{G}_{(a)}$*.*

La preuve de la Prop 0.0.3 est la première partie de la preuve de la Prop. 0.0.2. Dans le paragraphe suivant, nous réalisons la même analyse présentée ci-dessus pour le jeu $\mathcal{G}_{(b)}$.

Le Jeu de Sélection de Canal

Lorsque $K = 2$ et $S = 2$, le jeu $\mathcal{G}_{(b)}$ a quatre états possibles , i.e., $\left|\mathcal{P}^{(b)}\right| = 4$. Nous décrivons ces états et la fonction de potentiel correspondante dans la Fig. 5.

$Tx_1 \backslash Tx_2$	$\mathbf{p}_2 = (p_{\max}, 0)$	$\mathbf{p}_2 = (0, p_{\max})$
$\mathbf{p}_1 = (p_{\max}, 0)$	$\frac{1}{2}\log_2\left(\sigma^2 + p_{\max}(g_{11}+g_{21})\right) + \frac{1}{2}\log_2\left(\sigma^2\right)$	$\frac{1}{2}\log_2\left(\sigma^2 + p_{\max} g_{11}\right) + \frac{1}{2}\log_2\left(\sigma^2 + p_{\max} g_{22}\right)$
$\mathbf{p}_1 = (0, p_{\max})$	$\frac{1}{2}\log_2\left(\sigma^2 + p_{\max} g_{12}\right) + \frac{1}{2}\log_2\left(\sigma^2 + p_{\max} g_{21}\right)$	$\frac{1}{2}\log_2\left(\sigma^2 + p_{\max}(g_{12}+g_{22})\right) + \frac{1}{2}\log_2\left(\sigma^2\right)$

Figure 5: La fonction potentiel ϕ du jeu $\mathcal{G}_{(b)}$, avec $K = 2$ et $S = 2$. Le joueur 1 choisit les lignes et le joueur 2 choisit les colonnes.

Nous remarquons que la Déf. 0.0.3 implique que chacun de ces résultats peut être potentiellement un équilibre de Nash selon les réalisations des canaux $\{g_{i,j}\}_{\forall (i,j) \in \mathcal{K}\times\mathcal{P}}$, comme indiqué dans la proposition suivante.

Proposition 0.0.4 (Équilibre de Nash dans le jeu $\mathcal{G}_{(b)}$) *Notons par* $\mathbf{p}^* = (\mathbf{p}_1^*, \mathbf{p}_2^*) \in \mathcal{P}^{(b)}$ *un des équilibres du jeu* $\mathcal{G}_{(b)}$*. Ensuite, en fonction des gains de canaux* $\{g_{i,j}\}_{\forall (i,j) \in \mathcal{K}\times\mathcal{P}}$*, le vecteur* $\mathbf{p}^*$ *peut être écrit comme:*

- *Équilibre* 1*: lorsque* $\mathbf{g} \in \mathcal{A}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{1}{1+\mathrm{SNR} g_{22}}$ *and* $\frac{g_{21}}{g_{22}} \leqslant 1 + \mathrm{SNR} g_{11}\}$*,* $\mathbf{p}_1^* = (p_{\max}, 0)$ *et* $\mathbf{p}_2^* = (0, p_{\max})$*.*

- *Équilibre* 2*: lorsque* $\mathbf{g} \in \mathcal{A}_2 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 + \mathrm{SNR} g_{21}$ *et* $\frac{g_{21}}{g_{22}} \geqslant 1 + \mathrm{SNR} g_{11}\ \}$*,* $\mathbf{p}_1^* = (p_{\max}, 0)$ *and* $\mathbf{p}_2^* = (p_{\max}, 0)$*.*

- *Équilibre* 3*: lorsque* $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{A}_3\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1}{1+\mathrm{SNR} g_{22}}$ *et* $\frac{g_{21}}{g_{22}} \leqslant \frac{1}{1+\mathrm{SNR} g_{12}}\}$*,* $\mathbf{p}_1^* = (0, p_{\max})$ *et* $\mathbf{p}_2^* = (0, p_{\max})$*.*

- *Équilibre* 4*: lorsque* $\mathbf{g} \in \mathcal{A}_4 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant 1 + \mathrm{SNR} g_{12}$ *et* $\frac{g_{21}}{g_{22}} \geqslant \frac{1}{1+\mathrm{SNR} g_{12}}\ \}$*,* $\mathbf{p}_1^* = (0, p_{\max})$ *et* $\mathbf{p}_2^* = (p_{\max}, 0)$*.*

Proof: La preuve suit de la Déf. 0.0.3 et la Fig. 5. □

Dans la Fig. 6, nous avons tracé les différents types de équilibres de Nash en fonction des rapport de canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$. Nous soulignons comment les profils $\mathbf{p}^* = (p_{\max}, 0, 0, p_{\max})$ et $\mathbf{p}^+ = (0, p_{\max}, p_{\max}, 0)$ sont, les deux, des équilibres de Nash lorsque les réalisations de canaux satisfont la condition: $\mathbf{g} \in \mathcal{A}_5 = \mathcal{A}_1 \cap \mathcal{A}_4$, i.e., $\mathcal{A}_5 =$

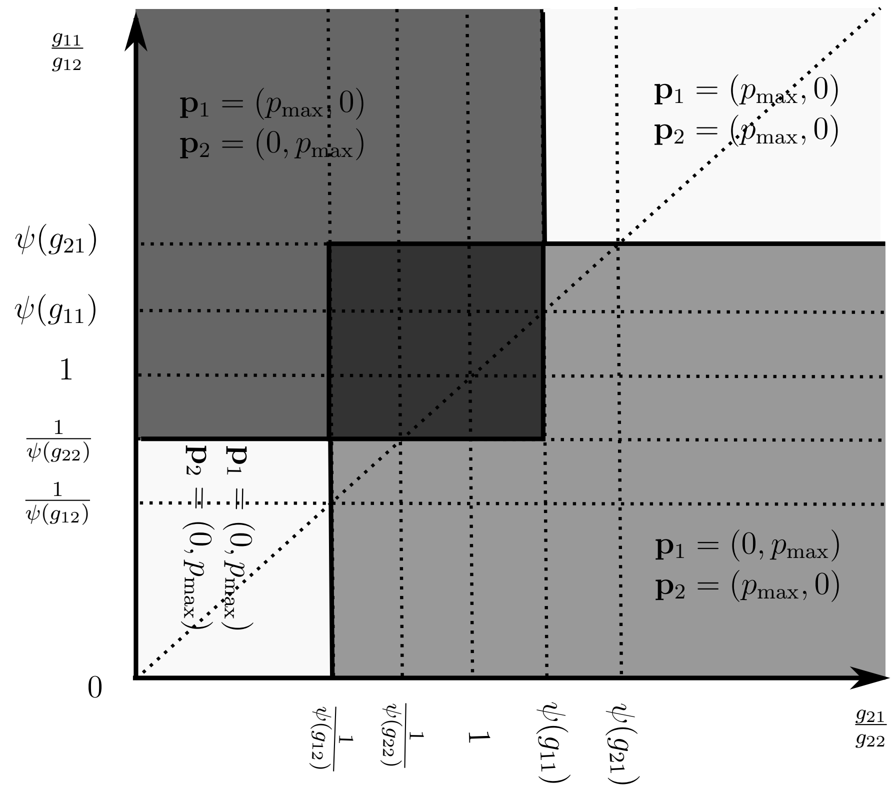

Figure 6: Équilibres de Nash comme fonction des rapports de canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$ pour le jeu à deux joueurs et deux canaux $\mathcal{G}_{(b)}$. La fonction $\psi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est définie comme suit: $\psi(x) = 1 + \mathrm{SNR}\, x$. Dans cet exemple, nous avons accordé arbitrairement $\psi(g_{11}) < \psi(g_{21})$.

$\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{1}{1+\mathrm{SNR}g_{22}} \leqslant \frac{g_{11}}{g_{12}} \leqslant 1 + \mathrm{SNR}g_{21}$ and $\frac{1}{1+\mathrm{SNR}g_{12}} \leqslant \frac{g_{21}}{g_{22}} \leqslant 1 + \mathrm{SNR}g_{11}\}$. Cela confirme que plusieurs équilibres peuvent exister dans le jeu $\mathcal{G}_{(b)}$ selon la réalisation exacte des canaux, comme indique la Prop: 3.4.2. Par ailleurs, on peut aussi observer qu'il pourrait exister un profil d'action correspondant à un équilibre de Nash qui n'est pas un optimum global de la fonction de potentiel (25) [113] (e.g., $\phi(\mathbf{p}^*) < \phi\left(\mathbf{p}_2^+\right)$). Maintenant, après le résultat dans [117], il peut être conclu que lorsqu'il existe deux NE en stratégies pures, il en existe un troisième en stratégies mixtes. Quand il existe un équilibre unique en stratégies pures, l'équilibre en stratégies mixtes coïncide avec celui en stratégies pures. Dans la suite, les performances atteintes par les deux émetteurs à l'équilibre de Nash dans les deux jeux de PA et CS sont comparées.

Un Paradoxe de Braess

Dans le jeu $\mathcal{G}_{(b)}$, l'ensemble d'actions du joueur k est un sous-ensemble de son ensemble d'actions dans le jeu $\mathcal{G}_{(a)}$, i.e., $\forall k \in \mathcal{K}$, $\mathcal{P}_k^{(b)} \subseteq \mathcal{P}_k^{(a)}$. On pourrait donc naïvement imaginer qu'un plus grand ensemble d'actions conduit à une meilleure performance globale, par exemple, une plus grande efficacité spectrale. Dans cette thèse, nous montrons que, contrairement à l'intuition, la réduction de l'ensemble d'actions de chaque joueur conduit à une meilleure performance globale. Cet effet (souvent associé à un paradoxe de Braess [15]) a été observé dans le canal à interférence en parallèle [90] et dans le canal à accès multiple en parallèle [75] pour le cas d'annulation successive d'interférence. Dans la suite, nous comparons l'efficacité spectrale globale obtenue dans les jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$ à l'équilibre de Nash.
Nous notons par $\mathbf{p}_k^{(\dagger,n)}$, l'unique équilibre du jeu $\mathcal{G}_{(a)}$, lorsque le vecteur

$$\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_n,$$

pour tout $n \in \{1\ldots,8\}$. Notons aussi par $\mathbf{p}^{(*,n)}$ un des équilibres du jeu $\mathcal{G}_{(b)}$ lorsque $(g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{A}_n$, pour tout $n \in \{1,\ldots 4\}$. Les ensembles $\mathcal{A}_n$ et $\mathcal{B}_n$ sont définis dans les Prop. 0.0.2 et Prop.0.0.4. Ansi, pour un niveau de RSB, SNR > 0, nous observons que $\forall n \in \{1,\ldots,4\}$, $\mathcal{A}_n \cap \mathcal{B}_n = \mathcal{B}_n$ et $\forall \mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_n$, l'égalité suivante est toujours satisfaite $\mathbf{p}_k^{(\dagger,n)} = \mathbf{p}_k^{(*,n)}$, ce qui implique la même performance dans les deux jeux. Toutefois, lorsque les équilibres des deux jeux sont différents, nous ne pouvons pas comparer facilement les utilités atteintes par chaque joueur, car elles dépendent des réalisations exactes des canaux. Heureusement, l'analyse simplifie largement en considérant soit un régime à faible RSB ou un régime à fort RSB.

Proposition 0.0.5 *Dans un régime à faible RSB, les deux jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$ possèdent un unique équilibre. Nous notons cet équilibre pour $k \in \mathcal{K}$ et $n_k \in \mathcal{S}$ comme,*

$$p_{k,n_k}^* = p_{\max} \mathbb{1}_{\left\{n_k = \arg\max\limits_{\ell \in \mathcal{S}} g_{k,\ell}\right\}} \quad (28)$$

$$p_{k,-n_k}^* = p_{\max} - p_{k,n_k}. \quad (29)$$

Proof: Voir l'annexe F □

De même, la comparaison des performances entre les jeux $\mathcal{G}_{(a)}$ et $\mathcal{G}_{(b)}$ pour le régime à fort RSB est présentée dans la proposition suivante.

Proposition 0.0.6 (Existence d'un Paradoxe de Braess) *Dans le régime à fort RSB, le jeu* $\mathcal{G}_{(a)}$ *a un unique équilibre pur que nous notons* $\mathbf{p}^{\dagger}$ *et le jeu* $\mathcal{G}_{(b)}$ *a deux équilibres purs que nous notons* $\mathbf{p}^{(*,1)}$ *et* $\mathbf{p}^{(*,4)}$. *Alors, il existe au moins un* $n \in \{1,4\}$ *et une valeur* $\mathrm{SNR}_0 > 0$, *telle que* $\forall \mathrm{SNR} \geqslant \mathrm{SNR}_0$

$$\sum_{k=1}^{2} u_k\big(\mathbf{p}^{(*,n)}\big) - \sum_{k=1}^{2} u_k(\mathbf{p}^{\dagger}) \geqslant \delta, \tag{30}$$

et $\delta \geqslant 0$.

Pour la preuve voir App. G. À partir de la Prop. 0.0.5 et la Prop. 0.0.6, nous pouvons conclure que dans le régime à faible RSB, les deux jeux $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ induisent la même efficacité spectrale globale. Au contraire, le jeu $\mathcal{G}_{(b)}$ induit toujours une efficacité spectrale égale ou supérieure à cela induite par le jeu $\mathcal{G}_{(b)}$ dans le régime à fort RSB. Ce résultat contre-intuitif implique un paradoxe de Braess, puisque $\mathcal{P}^{(b)} \subset \mathcal{P}^{(a)}$.

Conclusions

Dans cette thèse, il est clairement montré dans quelle mesure l'analyse des équilibres de Nash dans les canaux à accès multiple décentralisés en parallèle diffère de celle réalisée pour d'autres canaux comme les canaux MIMO Gaussien à interférence. La structure particulière des canaux à accès multiple décentralisés en parallèle (qui sont des modèles de canal important dans la pratique) sous l'hypothèse d'un décodage à unique utilisateur au récepteur conduit à la propriété des jeux de potentiel. Le jeu de sélection de canal avait été traité dans la littérature vaguement mais jamais avec le niveau de détail avec lequel il a été présenté dans cette thèse. En particulier, une interprétation en théorie des graphes est utilisée pour caractériser le nombre d'équilibre et la propriété des jeux de potentiel est exploitée pour développer des procédures d'apprentissage. Néanmoins, bien que tous ces résultats soient encourageants quant à la pertinence de la théorie des jeux pour l'analyse des problèmes d'allocation de puissance, d'importantes questions d'ordre pratique ont été délibérément ignorées. Par exemple, l'impact de l'estimation de canal n'est pas évalué dans cette thèse.

Apprentissage d'un Équilibre

Dans la section précédente, nous avons montré que la pertinence de la notion d'équilibre de Nash provient du fait qu'une fois qu'il est atteint, la configuration de transmission de chaque appareil radio est optimale par rapport aux configurations de transmission adoptées par tous les autres appareils. Ainsi, cette notion d'équilibre est clairement une solution désirée du point de vue de chaque appareil radio et semble

être pertinent dans les réseaux sans fil où ni la coordination ni la coopération entre tous les éléments du réseau n'est possible. Néanmoins, atteindre un équilibre dans un réseau décentralisé n'est pas une tâche facile. Comme une des principales contraintes, nous mettons en évidence le fait que les appareils radio ne sont pas capables d'observer ni la configuration de transmission (par exemple, le niveau de puissance d'émission ou canal sélectionné) des autres appareils, ni l'état global instantané du réseau, à savoir, les réalisations des canaux, les contraintes d'énergie et la qualité de service exigée par tous les appareils radio. Ainsi, le manque d'information de chaque appareil radio à tout instant donné devient naturellement une contrainte pour déterminer une règle de comportement qui lui permet d'atteindre un équilibre.
Dans cette perspective, un intérêt croissant a été observé dans la conception de règles de comportement pour permettre aux appareils radio d'atteindre une configuration d'équilibre comme un résultat d'une interaction de courte durée avec les autres appareils, semblable à un processus d'apprentissage [47]. Dans ce sens, la dynamique meilleure réponse (best response dynamics, BRD) [31] et le jeu fictif (Fictitious Play, FP) [16] ont été largement utilisés dans les communications sans fil [71, 73, 97, 98, 100] et il a été prouvé qu'ils convergent vers un équilibre dans des réseaux de certaines topologies.
La principale contrainte de la BRD, le FP et ses variantes est le fait que chaque appareil radio doit connaitre les configurations de transmission de tous les autres appareils, l'état du jeu instantané et en plus, ils doivent posséder une expression analytique de la fonction d'utilité, ce qui est clairement une condition très exigeante dans un scénario pratique. Dans certaines topologies de réseau et en fonction de la métrique de performance, cette condition peut être affaiblie et un simple message de diffusion à partir de chaque récepteur peut être suffisant pour mettre en oeuvre soit le BRD ou le FP [98]. Toutefois, le nombre de messages de signalisation requis pourrait être très élevé selon le nombre de dimensions du scénario, par exemple, le nombre de bandes de fréquences ou les antennes de transmission.
Des règles de comportement plus élaborées pour atteindre des équilibres sont basées sur l'apprentissage par renforcement (AR) [18, 95, 120]. Dans le AR, l'information requise par chaque appareil radio est tout simplement une observation de ses propres performances atteintes au moins chaque fois qu'il change sa configuration de transmission. Le principe de la RL est la suivante. Après avoir observé la valeur actuelle de son utilité, chaque appareil radio met à jour une distribution de probabilité sur l'ensemble de toutes ses configurations possibles (ou actions). À chaque mise à jour, la probabilité d'augmenter ou diminuer la probabilité d'utiliser une configuration en particulier dépend de l'utilité observée chaque fois qu'elle a été jouée. Dans le domaine des communications sans fil, cette idée a été utilisée dans certains scénarios, en particulier dans l'allocation dynamique des ressources radio [118, 123].
Les principaux avantages de l'AR à l'égard de la BRD et FP sont nombreux (à condition qu'elle converge vers NE). Par exemple, l'AR est moins exigeant en termes d'information: seulement une observation de la l'utilité atteinte à chaque étape de jeu est suffisant pour appliquer la règle de l'AR.
Cependant, en dehors de tous les avantages attrayants de l'AR, il a un inconvénient

essentiel: chaque observation de l'utilitaire est utilisé pour mettre à jour directement la distribution de probabilité de chaque appareil, sans le maintien d'une estimation de la performance atteinte avec chacune des configurations de transmission. Ce fait pourrait conduire le réseau à converger vers un état stationnaire qui n'est pas un équilibre. Nous disons stationnaire, dans le sens où aucun des appareil radio change sa configuration car il est incapable d'identifier d'autres configurations qui pourraient apporter une meilleure performance.
Motivés par cette observation, dans cette thèse, nous introduisons une sorte de règles de comportement qui sont connues dans le domaine des processus de décision de Markov comme des algorithmes d'acteur critique [43, 44, 107]. Ici, chaque appareil radio apprend simultanément la performance moyenne temporelle obtenue avec chacune de ses configurations de transmission et la distribution de probabilité d'équilibre. Cette estimation permet de résoudre le problème rencontré dans les règles de comportement fondées sur l'apprentissage par renforcement, où la convergence est observée, mais la configuration finale du réseau ne correspond pas à un véritable équilibre de Nash. En particulier, contrairement aux algorithmes de l'AR décrits ci-dessus, chaque fois que ces règles de comportement conduisent à une configuration de réseau stationnaire, cela correspond à un équilibre logit (EL), qui est en effet, un concept proche d'epsilon-équilibre de Nash.

Formulation du Jeu

Considérons le jeu $\mathcal{G}_{(b)}$ décrit dans la Sec. et supposons qu'il est joué de manière répété à l'infini. Chaque étape est considérée indépendante de toutes les étapes précédentes. À chaque étape n, chaque joueur k adopte une action, par exemple, un vecteur d'allocation de puissance $\mathbf{p}_k(n) \in \mathcal{A}_k$. A la fin de l'étape n, le joueur k observe une valeur numérique $\tilde{u}_k(n)$ de sa performance individuelle instantanée, c'est-à-dire, son efficacité spectrale $\tilde{u}_k(n) = u_k\left(\mathbf{h}(n), \mathbf{p}_k(n), \mathbf{p}_{-k}(n)\right)$. Nous soulignons que ces observations peuvent être bruitées [74]. Toutefois, ce cas n'est pas considéré dans cette thèse.

Nous noterons par $\theta_k(n)$ toute l'information rassemblée par le joueur k jusqu'à l'étape n, i.e.,

$$\theta_k(n) = \{(a_k(0), \tilde{u}_k(0)), \ldots, (a_k(n-1), \tilde{u}_k(n-1))\}. \tag{31}$$

À chaque étape de jeu, les émetteurs choisissent leurs vecteurs d'allocation de puissance respectifs en suivant une distribution de probabilité

$$\pi_k(n) = \left(\pi_{k,A_k^{(1)}}(n), \ldots, \pi_{k,A_k^{(N_k)}}(n)\right) \in \triangle\left(\mathcal{A}_k\right)$$

qui est construite en fonction de son histoire privée $\theta_k(n)$. Ici, $\forall n_k \in \{1, \ldots, N_k\}$, $\pi_{k,\mathbf{p}_k^{(n_k)}}(n)$ représente la probabilité que le joueur k joue l'action $\mathbf{p}_k^{(n_k)} \in \mathcal{A}_k$ à l'étape n, i.e.,

$$\pi_{k,A_k^{(n_k)}}(n) = \Pr\left(\mathbf{p}_k(n) = \mathbf{p}_k^{(n_k)}\right). \tag{32}$$

Cette probabilité, nommée en théorie de jeux comme stratégie, est dynamiquement mise à jour pour atteindre un équilibre. Dans la section suivante nous décrivons la notion d'équilibre que nous pouvons apprendre avec les utiles décrites dans cette section.

Équilibre Logit

Avant de fournir une définition formelle de l'équilibre logit, nous introduisons la notion de meilleure réponse logit.

Définition 0.0.5 (Meilleure Réponse logit) *Considérons le jeu $\mathcal{G}$ et notons la strategie du joueur k par $\pi_{-k} \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_{k-1}) \times \triangle(\mathcal{A}_{k+1}) \times \ldots \times \triangle(\mathcal{A}_K)$, avec $k \in \mathcal{K}$. Ansi, la meilleure réponse logit du joueur k, avec parametre $\gamma_k > 0$, est une distribution de probabilité $\beta_k^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})) \in \triangle(\mathcal{A}_k)$ telle que, $\beta_k^{(\gamma_k)} : \mathbb{R}^{N_k} \to \triangle(\mathcal{A}_k)$ est une fonction de logit,*

$$\beta_k^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot,\pi_{-k})) = \left(\beta_{k,A_k^{(1)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot,\pi_{-k})),\ldots,\beta_{k,A_k^{(N_k)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot,\pi_{-k}))\right)$$

et $\forall n_k \in \{1, \ldots, N_k\}$,

$$\beta_{k,A_k^{(n_k)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})) = \frac{\exp\left(\gamma_k \bar{u}_k(\mathbf{e}_k^{(n_k)}, \pi_{-k})\right)}{\sum_{m=1}^{N_k} \exp\left(\gamma_k \bar{u}_k(\mathbf{e}_k^{(m)}, \pi_{-k})\right)}. \tag{33}$$

À partir de la Déf. 0.0.5, il peut être conclu que, à chaque étape du jeu, chaque vecteur d'allocation de puissance d'un émetteur donné a une probabilité non nulle d'être joué, i.e., $\forall k \in \mathcal{K}$ and $\forall n_k \in \{1, \ldots, N_k\}$ et $\forall \gamma_k \in \mathbb{R}_+$, nous observons que $\beta_{k,A_k^{(n_k)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})) > 0$. Plus généralement, on peut affirmer que la meilleure réponse logit est représentée par une distribution de probabilité qui assigne des probabilités élevées aux vecteurs d'allocation de puissance associés à une haute efficacité spectrale et une faible probabilité aux vecteurs d'allocations de puissance associés à une faible efficacité spectrale
Enfin, nous signalons que contrairement à la meilleure réponse dans le cas général, la meilleure réponse logit du joueur k est unique pour toutes les stratégies que les autres joueurs pourraient adopter.
En utilisant la Déf. 0.0.5, nous introduisons la définition de l'équilibre logit:

Définition 0.0.6 (Équilibre Logit) *Considérons le jeu $\mathcal{G}$ et notons par $\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ un prolife de strategies en particulier. Ansi, le profil π^* est un équilibre logit avec parametre $\gamma = (\gamma_1, \ldots, \gamma_K)$, si pour tout $k \in \mathcal{K}$,*

$$\pi_k^* = \beta_k^{(\gamma_k)}\left(\bar{u}_k\left(\mathbf{e}_1^{(N_k)}, \pi_{-k}^*\right), \ldots, \bar{u}_k\left(\mathbf{e}_{N_k}^{(N_k)}, \pi_{-k}^*\right)\right). \tag{34}$$

À l'équilibre logit, puisque toutes les actions sont jouées avec une probabilité non nulle, les actions jouées par le joueur k dans certaines étapes ne maximisent pas sa performance instantanée u_k, ce qui affecte négativement la performance à long terme $\bar{u}_K$. Dans cette thèse, nous allons montrer que la perte maximale en performance du joueur k n'est pas plus élevée que $\frac{1}{\gamma_k}\ln(N_k)$, ce qui confirme que l'équilibre logit est dans la classe de ϵ-équilibre décrit dans la Déf. 4.2.1. Les aspects d'existence et unicité de l'équilibre logit sont aussi traités dans cette thèse dans le chapitre 4.

Apprentisage d'un Èquilibre Logit

Dans cette section, nous concevons des règles de comportement tels qu'étant donné l'informations recueillie par le joueur k à chaque étape n, i.e., les ensembles $\{\theta_k(n)\}_{n>0}$, tous les joueurs sont capables de produire une séquence $\{\pi_k(n)\}_{n>0}$, telle que, $\lim_{n\to\infty} ||\pi_k\,(n) - \pi_k^*|| = 0$, où

$$\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle\,(\mathcal{A}_1) \times \ldots \times \triangle\,(\mathcal{A}_K)$$

est un équilibre logit du jeu $\mathcal{G}$ (Déf. 0.0.6).
Une remarque importante à partir de la Déf. 0.0.6 est le fait que nous supposons que tous les joueurs sont capables de construire leurs meilleurs réponses logit. Néanmoins, la construction d'une telle distribution de probabilité exige à chaque joueur k de connaitre le vecteur $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$. Ainsi, les dispositifs radio doivent estimer leur vecteur correspondant $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ à chaque étape n du jeu en fonction de leur l'histoire actuelle $\theta_k(n)$ afin de générer leur meilleure réponse logit.
Nous notons par

$$\hat{\mathbf{u}}_k(n) = \left(\hat{u}_{k,A_k^{(1)}}(n), \ldots, \hat{u}_{k,A_k^{(N_k)}}(n)\right) \tag{35}$$

la version estimée par le joueur k du vecteur $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ à l'étape n du jeu. Dans la suite, nous présentons un résultat initialement introduit dans [82], qui permet aux appareils radio d'estimer simultanément leurs vecteurs $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ et de déterminer la distribution de probabilité $\pi_k(n)$, avec laquelle le joueur choisit le vecteur d'allocation de puissance $\mathbf{p}_k(n)$.

Théorème 0.0.5 (Règle de Comportement) *Considérons le jeu $\mathcal{G}$ et assumons que pour tout $k \in \mathcal{K}$ et pour tout $n_k \in \{1, \ldots, N_k\}$ nous définissons pour tout $n \in \mathbb{N}$,*

$$\begin{cases} \hat{u}_{k,\mathbf{p}_k^{(n_k)}}(n) & = & \hat{u}_{k,\mathbf{p}_k^{(n_k)}}(n-1) + \\ & & \alpha_k(n) \dfrac{\mathbb{1}_{\left\{\mathbf{p}_k(n-1)=\mathbf{p}_k^{(n_k)}\right\}}}{\pi_{k,\mathbf{p}_k^{(n_k)}}(n)} \left(\tilde{u}_k(n-1) - \hat{u}_{k,\mathbf{p}_k^{(n_k)}}(n-1)\right), \\ \pi_{k,\mathbf{p}_k^{(n_k)}}(n) & = & \pi_{k,\mathbf{p}_k^{(n_k)}}(n-1) + \\ & & \lambda_k(n) \left(\beta^{(\gamma_k)}_{k,\mathbf{p}_k^{(n_k)}}\left(\hat{\mathbf{u}}_k(n)\right) - \pi_{k,\mathbf{p}_k^{(n_k)}}(n-1)\right), \end{cases} \tag{36}$$

où, $\mathbf{p}_k(0) \in \mathcal{A}_k$, $\hat{\mathbf{u}}_k(0) \in \mathbb{R}^{N_k}$ and $\pi_k(0) \in \triangle\,(\mathcal{A}_k)$ sont les valuers initiales. Considérons aussi les hypotheses suivantes tel que pour tout $(j,k) \in \mathcal{K}^2$, le taux *d'apprentissage α_k et λ_j satisfont les contraintes :*

$$
\begin{aligned}
(B0) \quad & \lim_{T\to\infty}\sum_{n=0}^{T}\alpha_k(n) = +\infty \quad \textit{and} \quad \lim_{T\to\infty}\sum_{n=0}^{T}\alpha_k(n)^2 < +\infty,\\
(B1) \quad & \lim_{T\to\infty}\sum_{n=0}^{T}\lambda_k(n) = +\infty \quad \textit{and} \quad \lim_{T\to\infty}\sum_{n=0}^{T}\lambda_k(n)^2 < +\infty,
\end{aligned}
$$

et,

$$(B2) \quad \lim_{n\to\infty}\frac{\lambda_j(n)}{\alpha_k(n)} = 0.$$

Ansi, si l'ensemble des algorithmes couplés (36) d'approximation stochastique convergent, il résulte que :

$$
\begin{aligned}
\lim_{n\to\infty}\pi_k(n) &= \pi_k^*, && (37)\\
\lim_{n\to\infty}\hat{u}_{k,\mathbf{p}_k^{(n_k)}}(n) &= \bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*), && (38)
\end{aligned}
$$

où $\pi_k^* \in \triangle\left(\mathcal{A}_k\right)$ satisfait que

$$\pi_k^* = \beta_k^{(\gamma_k)}\left(\bar{u}_k\left(\mathbf{e}_1^{(N_k)},\pi_{-k}^*\right),\ldots,\bar{u}_k\left(\mathbf{e}_{N_k}^{(N_k)},\pi_{-k}^*\right)\right). \quad (39)$$

La preuve du théorème 0.0.5 est présentée dans le cas le plus général du jeu $\mathcal{G}$ dans le chapitre 4. Cette preuve utilise des résultats antérieurs sur des approximations stochastiques [12,48]. Dans cette thèse, nous prouvons que la règle de comportement dans le théorème 0.0.5 converge dans plusieurs classes de jeux, par exemple, des jeux de potentiels [64]. Cependant, en général, pour un jeu que n'appartient pas une de ces classes de jeux, la convergence vers un équilibre doit être prouvée.

Conclusions

Dans cette thèse, une dynamique d'apprentissage bien adaptée aux contraintes réelles des systèmes de communications a été introduite afin de permettre aux réseaux sans fil entièrement décentralisés d'atteindre un équilibre. Par exemple, il a été supposé que la seule information qu'un dispositif radio peut obtenir à partir du réseau est une mesure de sa performance instantanée, et, chaque appareil radio est complètement ignorant de l'existence de tous les autres appareils radio. Dans ces conditions, nos dynamiques d'apprentissage permettent que chaque appareil radio soit capable d'apprendre simultanément la stratégie et la performance atteinte avec chacune de ses actions à l'équilibre. Ici, nous avons utilisé des outils récents d'approximations stochastiques pour étudier la convergence d'une telle dynamique. En particulier, nous montrons qu'il existe plusieurs classes de jeux où ces dynamiques d'apprentissage convergent toujours vers un équilibre Logit.

Équilibrie de Satisfactions

Dans le contexte de réseaux auto-configurables, un équilibre de Nash (NE) est un état du réseau auquel les dispositifs radios ne peuvent pas améliorer leur qualité de service en changeant unilatéralement leur configuration de transmission. À l'équilibre de Nash, chaque appareil radio atteint le plus haut niveau possible de qualité de service étant donné les configurations de transmission de ses homologues. Cependant, d'un point de vue pratique, n'importe quel élément du réseau pourrait être plus intéressé à garantir un niveau minimum de qualité de service au lieu du niveau le plus élevé possible. Nous avons plusieurs raisons pour justifier cette affirmation. Tout d'abord, une communication devient possible lorsque certaines conditions spécifiques sont satisfaites, par exemple, un niveau minimum du RSBI, un délai minimum, *etc.* En plus, des niveaux plus élevés de qualité de service impliquent souvent des efforts plus importants, par exemple, des niveaux plus élevés de puissance de transmission, de traitement du signal plus complexe, *etc.* Finalement, nous signalons qu'en augmentant la qualité de service pour une communication particulière nous diminuons significativement la qualité des autres communications. Ce raisonnement implique que, en termes pratiques, le concept d'équilibre de Nash n'est pas la notion d'équilibre la plus adaptée pour modéliser un réseau décentralisé de communications. En présence de contraintes, en termes des niveaux de qualité de service minimum, une solution plus adaptée pour les réseaux décentralisés de communications est le concept d'équilibre introduit par Debreu dans [24] et aujourd'hui connu sous le nom de équilibre de Nash généralisé (GNE). Dans le contexte de réseaux décentralisés, un GNE est un état dans lequel les émetteurs satisfont leurs contraintes de qualité de service et leur performance ne peut pas être améliorée par des déviations unilatérales (comme dans le NE). Néanmoins, selon les paramètres utilisés pour mesurer la qualité de service et de la topologie du réseau, le GEN pourrait ne pas exister [46]. Dans le cas où, au moins un GNE existe, un émetteur finit toujours par atteindre le niveau plus élevé de qualité de service possible, ce qui est souvent coûteux comme mentionné ci-dessus. Dans le cas le plus général, on peut considérer que les émetteurs visent à satisfaire uniquement leurs contraintes au lieu de considérer qu'ils visent à maximiser leur performance au même temps qu'ils visent à satisfaire leurs contraintes. Ce raisonnement conduit à un autre type de concept d'équilibre: tout état d'un jeu donné où tous les joueurs satisfont leurs propres contraintes est un équilibre. Récemment, Ross *et al.* [93] ont formalisé ce concept pour un type particulier de contraintes. Un tel équilibre est appelé équilibre de satisfaction (SE) par les auteurs de [93]. Dans notre scénario, un SE représente tout état du réseau où tous les émetteurs satisfont leurs contraintes de QoS, indépendamment de leur performance atteinte.

Dans ce cadre, les contributions présentées dans cette thèse sont les suivants:

- Les notions de SF et SE sont formalisées dans le contexte des stratégies pures (pure strategies, PS) et des stratégies mixtes (mixed strategies, MS) pour des jeux finis. Quelques conditions pour l'existence de la SE en PS et MS sont établies.

- Nous introduisons la notion d'epsilon équilibre de satisfaction (ϵ-SE), qui est tout simplement une stratégie mixte qui permet à tous les joueurs d'être satisfait avec une probabilité d'au moins $1 - \epsilon$. Ce concept d'équilibre est moins restrictif en termes d'existence que le SE et donc, très importante pour le domaine des réseaux de communications décentralisés.

- Un raffinement de la notion de SE à laquelle nous nous référons comme équilibre efficace de satisfaction (efficient satisfaction equilibrium, ESE) est présenté comme un mécanisme de sélection d'équilibre impliquant l'idée d'effort de satisfaction.

- Un algorithme d'apprentissage simple à mettre en oeuvre pour atteindre le SE, basé sur les algorithmes proposés dans [94], est présenté.

Jeux Sous Forme de Satisfaction et Équilibre de Satisfaction

Dans cette section, nous présentons une nouvelle formulation de jeu où, contrairement aux formulations existantes (par exemple, la forme normale [66] et la forme normale avec des contraintes sous les ensembles d'actions [24]), l'idée d'optimisation des performances, i.e., maximisation de l'utilité ou la minimisation des coûts, n'existe pas. Dans notre formulation, à laquelle nous nous référons comme forme de satisfaction, le but des joueurs est d'adopter l'une des actions qui leur permet de satisfaire ses contraintes individuelles étant données les actions adoptées par tous les autres joueurs. Selon cette formulation de jeu, nous introduisons le concept d'équilibre de satisfaction.

Jeux sous Forme de Satisfaction

En général, un jeu sous la forme de satisfaction peut être décrit par les trois éléments suivants:

$$\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}. \tag{40}$$

Ici, l'ensemble $\mathcal{K} = \{1, \ldots, K\}$ répresente l'ensemble de joueurs et l'ensemble $\mathcal{A}_k = \{A_k^{(1)}, \ldots, A_k^{(N_k)}\}$ répresente l'ensemble de N_k actions disponibles pour un transmetteur donné k. Un profil d'action est un vecteur $\mathbf{a} = (a_1, \ldots, a_K) \in \mathcal{A}$, où,

$$\mathcal{A} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_K. \tag{41}$$

Dans cette analyse, les ensembles $\mathcal{K}$ and $\{\mathcal{A}_k\}_{k\in\mathcal{K}}$ sont supposés finis et non vides. Nous notons par $\mathbf{a}_{-k} = (a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_K) \in \mathcal{A}_{-k}$,où,

$$\mathcal{A}_{-k} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_{k-1} \times \mathcal{A}_{k+1} \times \ldots, \times\mathcal{A}_K. \tag{42}$$

La correspondance $f_k : \mathcal{A}_{-k} \to 2^{\mathcal{A}_k}$ détermine l'ensemble des actions du joueur k qui permet sa satisfaction en prenant en compte les actions jouées par tous les autres joueurs. Ici, la notation $2^{\mathcal{A}_k}$ désigne l'ensemble de tous les sous-ensembles possibles de l'ensemble $\mathcal{A}_k$, y compris $\mathcal{A}_k$. Nous soulignons que $2^{\mathcal{A}_k}$ inclut également l'ensemble vide, qui modélise le cas où un joueur se retrouve sans une action qui lui permet de satisfaire ses contraintes individuelles étant données les actions des autres joueurs. Souvent, c'est une contrainte forte en mathématiques et donc, dans

certaines sections de cette thèse, nous supposons qu'aucune des correspondances f_k n'est vide.
En général, un résultat important d'un jeu sous forme de satisfaction est celui où tous les joueurs sont satisfaits. Nous nous référons à ce résultat comme l'équilibre de satisfaction (SE).

Définition 0.0.7 (Équilibre de Satisfaction [83]) *Un profil d'action* $\mathbf{a}^+$ *est un équilibre de satisfaction du jeu* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$*, si*

$$\forall k \in \mathcal{K}, \quad a_k^+ \in f_k\left(\mathbf{a}_{-k}^+\right). \tag{43}$$

Nous remarquons que sous cette formulation, les résultats où tous les joueurs sont satisfaits est naturellement un équilibre. Ici, puisque le but de chaque joueur est d'être satisfait, aucun d'entre eux a un intérêt particulier à changer son action actuelle. Une remarque importante ici est que, dans cette formulation, les joueurs sont supposés n'est pas se préocuper du fait que les autres joueurs puissent ou non satisfaire leurs contraintes individuelles. Une analyse intéressante de l'impact de cette hypothèse dans la définition de l'équilibre Nash peut être trouvée dans [2].
Dans ce contexte, quand les réseaux décentralisés sont modélisés en utilisant la forme de satisfaction, les appareils radio sont indifférents au fait qu'il pourrait exister une autre configuration de transmission avec laquelle une meilleure performance peut être atteinte. Ici, lorsque chaque appareil radio est en mesure de satisfaire ses conditions individuelles de QoS, il n'a aucune incitation à modifier unilatéralement sa configuration de transmission ou réception.

Existence et Unicité de l'Équilibre de Satisfaction

Dans cette section, nous étudions l'existence et l'unicité d'un équilibre de satisfaction dans les jeux sous forme de satisfaction et dans son extension correspondante en stratégies mixtes. Une attention particulière est donnée à l'existence du ϵ-SE dans le cas où il n'existe pas un SE ni en stratégies pures ni en stratégies mixtes.

Existence du SE en Stratégies Pures

Afin d'étudier l'existence du SE dans le jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$, notons par $F : \mathcal{A} \to 2^{\mathcal{A}}$ la correspondance définie comme suit:

$$F(\mathbf{a}) = f_1\left(\mathbf{a}_{-1}\right) \times \ldots \times f_K\left(\mathbf{a}_{-K}\right). \tag{44}$$

Alors, un SE existe si et seulement si

$$\exists \mathbf{a} \in \mathcal{A} : \quad \mathbf{a} \in F(\mathbf{a}). \tag{45}$$

Nous soulignons que cette formulation nous permet d'utiliser les théorèmes de point fixe (FP) pour déterminer les conditions suffisantes pour l'existence d'au moins un SE. Par exemple, on peut compter sur le théorème de point fixe de Knaster et Tarski [42] pour énoncer le théorème suivant.

Théorème 0.0.6 (Existence d'un SE dans les jeux finis) *Considérons le jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ et supposons que l'ensemble $\mathcal{A}$ a une relation binaire notée par $\preceq$. Assumons aussi que: (i) $V = \langle \mathcal{A}, \preceq \rangle$ est un treillis complet;*
(ii) $F(\mathbf{a})$ est non vide pour tout $\mathbf{a} \in \mathcal{A}$;
(iii) la correspondence F (44) satisfait que $\forall (\mathbf{a}, \mathbf{a}') \in \mathcal{A}^2$, avec $\mathbf{a} \preceq \mathbf{a}'$,

$$\forall (\mathbf{b}, \mathbf{b}') \in F(\mathbf{a}) \times F(\mathbf{a}'), \quad \mathbf{b} \preceq \mathbf{b}'. \tag{46}$$

Puis le jeu a au moins un SE en stratégies pures.

Notez que le théorème 0.0.6 exige que pour tout $\mathbf{a} \in \mathcal{A}$, l'ensemble $F(\mathbf{a})$ est non vide, i.e.,

$$\forall k \in \mathcal{K} \text{ and } \forall \mathbf{a}_{-k} \in \mathcal{A}_{-k},\ \exists a_k \in \mathcal{A}_k : \ a_k \in f_k(\mathbf{a}_{-k}). \tag{47}$$

Dans certains cas, cette condition peut paraître restrictive. Toutefois, dans le contexte général des communications sans fil, quand un appareil radio n'est pas capable de satisfaire ses contraintes de QoS, l'action par défaut est tout simplement la mise en veille. Cela pourrait impliquer l'existence d'une action " ne rien faire (do nothing, DN)" qui pourrait être considérée pour éviter que $f_k(\mathbf{a}_{-k})$ soit vide, lorsque cela est nécessaire. L'interprétation de l'existence de l'action DN dépend fortement du scénario. Par exemple, dans le cas des jeux d'allocation de puissance, une telle action peut être le vecteur nul, c'est-à-dire, l'équivalent à une puissance d'émission zéro.
En général, il est difficile de fournir les conditions nécessaires pour observer un SE unique pour un ensemble général de correspondances $\{f_k\}_{k\in\mathcal{K}}$. Comme nous le verrons dans la suite de cette thèse, l'ensemble des SE n'est souvent pas unitaire dans les jeux qui modélisent des réseaux décentralisés. Dans ce contexte, nous avons fourni un mécanisme de sélection d'équilibres.

Apprentissage de l'Équilibre de Satisfaction

Dans cette section, nous étudions une règle de comportement qui permet aux appareils radio d'apprendre l'équilibre de satisfaction d'une manière totalement décentralisée. Ici, l'hypothèse à souligner est que les joueurs n'ont pas besoin d'observer la valeur exacte de l'utilité instantanée atteinte, à savoir, le taux de transmission, l'efficacité énergétique, etc, mais seulement de savoir s'ils sont satisfaits ou non à chaque étape du processus d'apprentissage. Cela implique seulement en échange de messages de longueur 1-bit entre les correspondants émetteur-récepteur. Dans la suite, nous formulons le problème d'apprentissage correspondant et, plus tard, nous introduisons des règles comportementales qui permettent aux joueurs d'apprendre la SE.

Formulation du Problème d'Apprentissage

Nous décrivons le processus d'apprentissage d'un SE en termes d'éléments du jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ comme suit. Supposons que le temps est divisé en étapes notées par l'indice $n \in \mathbb{N}$. Chaque étape se termine quand chaque joueur a joué une fois. Nous notons les actions prises par le joueur k à l'étape n par $a_k(n)$.

À chaque intervalle de n, le joueur k observe s'il est satisfait ou non, c'est-à-dire, il observe une variable binaire

$$\tilde{v}_k(n) = \mathbb{1}_{\{a_k(n) \in f_k(\mathbf{a}_{-k}(n))\}}. \tag{48}$$

Notre intention est d'apprendre au moins un SE en laissant les joueurs interagir en suivant des règles de comportement particulières. Nous disons que les joueurs apprennent un équilibre de satisfaction en stratégies pures si, après un nombre fini d'intervalles de temps donnée, tous les joueurs ont choisi une action qui atteint sa satisfaction, et donc, aucune mise à jour des actions n'a lieu à partir de ce moment.

Apprentissage du SE en Strategies Pures

Avant de présenter la règle de comportement qui permet aux joueurs d'atteindre l'un des équilibres du jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$, nous déclarons l'hypothèse suivante:

(i) Le jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ a au moins un SE en stratégies pures.

(ii) Pour tout $k \in \mathcal{K}$ et pour tout $\mathbf{a}_{-k} \in \mathcal{A}_{-k}$, l'ensemble $f_k(\mathbf{a}_{-k})$ n'est pas vide.

(iii) Les ensembles $\mathcal{K}$ et $\{\mathcal{A}_k\}_{k\in\mathcal{K}}$ sont finis.

La première hypothèse garantit que le problème d'apprentissage SE est bien posé, à savoir, les appareils radio sont assignés une tâche réalisable. La seconde hypothèse fait référence au fait que, chaque appareil radio est toujours en mesure de trouver une configuration d'émission/réception avec laquelle il peut être considéré comme satisfait étant données les configurations d'émission/réception de tous les autres appareils radio. Cette hypothèse peut sembler restrictive, mais ce n'est pas nécessairement le cas, voir la discussion sur l'action "ne rien faire" dans la Sec. . La troisième hypothèse est considérée afin de s'assurer que notre algorithme est capable de converger dans un temps fini.
Sous la condition que toutes les hypothèses sont vraies pour un jeu en particulier, chaque joueur choisit ses actions comme suit. La première action du joueur k, notée $a_k(0)$, est prise en considérant une distribution de probabilité arbitraire $\hat{\pi}_k(0) \in \triangle(\mathcal{A}_k)$. Souvent, cette probabilité $\hat{\pi}_k(0)$ est la distribution uniforme. À l'étape $n > 0$, le joueur k change son action si et seulement si il n'est pas satisfait, c'est-à-dire, si $\tilde{v}_k(n-1) = 0$. Dans ce cas, la prochaine action est choisi en suivant une distribution de probabilité $\hat{\pi}_k(n)$ à laquelle nous nous référons comme la distribution de probabilité de l'exploration. Si le joueur k est satisfait à l'étape n, c'est-à-dire, $\tilde{v}_k(n-1) = 1$, il continue à jouer la même action. Ainsi, on peut écrire que,

$$a_k(n) = \begin{cases} a_k(n-1) & \text{if} \quad \tilde{v}_k(n-1) = 1 \\ a_k(n) \sim \hat{\pi}_k(n) & \text{if} \quad \tilde{v}_k(n-1) = 0 \end{cases}. \tag{49}$$

La règle de comportement (49) est basée sur la proposition présentée en [94]. Nous remarquons que dans cette formulation, la seule amélioration possible est sur la conception de $\hat{\pi}_k(n)$ et son évolution au fil du temps. Toutefois, nous avons laissé cette question hors de la portée de cette thèse et aucune distribution de probabilité particulière est supposé. Ici nous utilisons la distribution uniforme. Nous formalisons la règle de comportement (49) dans l'algorithme présenté dans Alg. 1, dans le Chapitre 5. En ce qui concerne la convergence de cette règle de comportement, nous fournissons la proposition suivante.

Proposition 0.0.7 *La règle de comportement* (49) *avec une distribution de probabilité* $\pi_k = \left(\pi_{k,A_k^{(1)}}, \ldots, \pi_{k,A_k^{(N_k)}}\right) \in \triangle(\mathcal{A}_k)$*, avec* $k \in \mathcal{K}$*, converge vers un SE du jeu* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *en temps fini si pour tout* $k \in \mathcal{K}$ *et pour tout* $n_k \in \{1, \ldots, N_k\}$,

$$\pi_{k,A_k^{(n_k)}}(n) > 0, \tag{50}$$

à chaque étape $n \in \mathbb{N}$*, et sous la condition que les hypothèses* (i)*,* (ii) *and* (iii) *sont toujours vraies.*

Les Strategies de Clipping et le SE

La règle de comportement (49) converge vers un SE en stratégies pures en un temps fini. Toutefois, dans les scénarios du système réel, il est souvent observé qu'il pourrait y exister une action d'un joueur donné, qui atteint la satisfaction de ce joueur en particulier indépendamment des actions adoptées par tous les autres joueurs. Nous nous référons à ce type d'actions comment *actions de clipping* (actions de clippings, CA) [83].

Définition 0.0.8 (Clipping Action) *Nous disons que un joueur* k *dans le jeu* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *a une* action de clipping $a_k \in \mathcal{A}_k$ *si*

$$\forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}, \quad a_k \in f_k\left(\mathbf{a}_{-k}\right). \tag{51}$$

Comme le montre la proposition suivante, l'existence d'une *action de clippings* dans le jeu $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ pourrait contraindre la convergence de la règle de comportement dans (49).

Proposition 0.0.8 *Considérons le jeu* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *sous la forme de satisfaction. Assumons l'existence d'au moins une* actions de clippings *que nous noterons par* $a_k^* \in \mathcal{A}_k$ *pour le joueur* k*, avec* $k \in \mathcal{K}$*. S'il existe un joueur* $j \in \mathcal{K} \setminus \{k\}$ *pour lequel* $f_j\left(a_k^*, \mathbf{a}_{-\{j,k\}}\right) = \emptyset$*,* $\forall \mathbf{a}_{-\{j,k\}} \in \prod_{i\in\mathcal{K}\setminus\{j,k\}} \mathcal{A}_i$*, la règle de comportement dans* (49) *ne converge pas vers un SE, avec une probabilité strictement positive.*

Il existe des alternatives simples qui peuvent être utilisés pour résoudre ce problème de convergence. Par exemple, la règle de comportement dans (49) peut être modifié de telle sorte que un joueur change son action en cours (en utilisant une distribution de probabilité donnée sur les actions), même si elle est satisfaite lorsque ce joueur voit les autres joueurs pas satisfait pendant une longue période. Néanmoins, dans ce cas, les joueurs auraient besoin d'avoir plus de 1-bit de rétroalimentation en vue de détecter la non-satisfaction des autres. Par exemple, les valeurs instantanées de la métrique de performance. Cette approche peut être comparée avec l'idée d'expérimentation epsilon discuté dans [120].

Conclusions

La formulation sous forme de satisfaction (SF) et la notion d'équilibre de satisfaction (SE) introduits dans cette thèse sont, comme nous le montrons dans les chapitres qui suivent, parfaitement adaptés pour modéliser le problème de l'approvisionnement décentralisé de qualité de service. Lorsque le réseau est dans un état d'équilibre de satisfaction, tous les joueurs sont satisfaits. Au contraire, lorsque le problème de QoS est modélisé par un jeux sous forme normale classique ou sous forme normale avec des spaces d'actions contraints, il possible d'observer des équilibres où tous les joueurs ne sont pas satisfait, même quand il existe des profils d'action qui permettent la satisfaction simultanée des tous les joueurs. La notion de SE a été formalisée dans le cadre de stratégies pures et mixtes et de son existence et unicité a été étudiée. En particulier, lorsque aucune SE existe ni en stratégies pures, ni mixtes, les conditions nécessaires et suffisantes pour l'existence d'un epsilon-SE a été présenté. Enfin, une dynamique d'apprentissage a été proposé pour atteindre le SE. En particulier, on remarque que pour implementer tels algorithmes il suffit seulement de permettre des échanges de messages de maximum 1-bit entre les correspondants émetteur-récepteur. Néanmoins, la convergence reste toujours conditionnée. Cela suggère que la conception d'algorithmes tels qu'au moins un SE est atteint dans en un temps fini et d'une façon totalement distribuée reste étant un problème ouvert.

Perspectives

Dans la suite, nous décrivons les perspectives que nous avons identifiées pour donner continuité aux travaux présentés dans cette thèse.

En ce qui concerne les extensions de notre schéma d'alignement opportuniste d'interférence, nous rappelons que notre solution est proposée pour un contexte où il y a que deux liens MIMO uniquement. C'est-à-dire, un unique lien primaire et un unique lien secondaire. Le cas où il y a plusieurs liens opportunistes et/ou plusieurs liens primaires reste à être étudié en détail. Par ailleurs, d'autres auteurs ont récemment montré que la notion d'alignement opportuniste d'interférence peut être utilisée dans autres topologies de réseau. Par exemple, sur le lien montant de systèmes de communication cellulaires, ce qui élargisse énormément les applications de cette technique. Ici, nous n'avons traité que le cas du canal à interférence. Un autre point également important regarde les hypothèses d'information. Ces hypothèses pourraient être assouplies pour rendre l'approche proposée plus pratique. Cette remarque concerne les hypothèses sur l'état du canal (channel state information, CSI) et aussi des hypothèses de comportement du système primaire. Dans le cas du CSI, la manque d'information dans le système secondaire peut être abordé en utilisant des algorithmes d'apprentissage. Toutefois, une procédure d'apprentisage implique que le système primaire doit être tolérant à une certaine quantité d'interférence du système opportuniste au cours de la période d'apprentissage. Dans le cas des hypothèses de comportement, il a été supposé que le régime de précodage utilisé par l'émetteur principal atteint la capacité du canal, ce qui permet à l'émetteur secondaire de prédire comment l'émetteur secondaire va exploiter ses ressources spatiales. Cette hypothèse comportementale pourrait être assouplie, mais certains mécanismes de détection doivent être conçus pour savoir quels modes spatiaux peuvent être utilisés efficacement par l'émetteur secondaire. Cette dernière idée pourrait être une extension très intéressante du schéma proposé.

Dans le contexte des études des équilibres de Nash présentés dans cette thèse, nous soulignons que l'intérêt sur ces équilibres repose sur le fait qu'ils permetent de fournir une prévision de la performance du réseau. Toutefois, nous avons montré que cet analyse dépend fortement de la topologie du réseau. Ici, un cadre unifié pour l'analyse des réseaux décentralisés indépendamment de leur topologies reste manquante. L'importance de la recherche dans ce sens, repose sur le fait que la topologie des ce sort de réseau est en constante évolution. Ainsi, une analyse telle que celle présentée dans cette thèse, est limitée aux hypothèses sur le temps de cohérence des canaux. Idéalement, un cadre général pour cet analyse doit prendre en considération ces faits. Nous accordons aussi une attention particulière au fait que les appareils radio fonctionnent que pendant la période dont un besoin de communication existe. En conséquence, le nombre d'émetteurs actifs est aussi variable. Une autre direction de recherche est de considérer qu'il existe un sous-ensemble des appareils radio qui ne sont pas très fiables. Par exemple, nous pouvons considérer qu'il existent des dispositifs radio visant à rompre les communications, soit parce qu'il est dans son propre intérêt ou tout simplement parce que les éléments extérieurs sont conçus pour attaquer le réseau. Ce comportement malveillant n'a pas été pris en compte dans cette analyse.

L'apprentissage des équilibres est l'une des lignes les plus intéressantes pour donner une continuité aux travaux présentés dans cette thèse. Comme nous l'avons mentionné précédemment, les différents concepts d'équilibre, par exemple l'équilibre de Nash, l'équilibre corrélé, l'équilibre de satisfaction nous ont permit de fournir une estimation de la performance du réseau. Toutefois, atteindre cet équilibre dans un réseau entièrement décentralisé reste une question ouverte. Comme nous l'avons montré dans cette thèse, dans certaines topologies de réseau, des règles de comportement très simples, par exemple, la dynamique de meilleure réponse (best response dinamics, BRD) ou un jeu fictif (fictitious play, FP), conduisent à un équilibre. Cependant, la convergence ou la non-convergence de ces algorithmes est fortement dépendante de la topologie de réseau. Par exemple, lorsque les appareil radio visent à maximiser leurs taux de transmission, la BRD et le FP convergent vers un NE dans le canal à accès multiple parallèle, contrairement au cas du canal à interférence parallèle, où une telle convergence n'est pas assurée. En général, une règle de comportement qui permet à tous les appareils radio d'atteindre un équilibre en temps fini, indépendamment de la topologie, n'existe pas. Souvent, les algorithmes atteignent des performances epsilon-près de l'équilibre après certain nombre d'itérations. Toutefois, en fonction de l'application, la performance globale du réseau pourrait être très sensible à ce temps d'apprentissage. Ainsi, la conception d'algorithmes qui permettent d'atteindre des performances d'équilibre dans une courte période de temps avec un minimum d'information reste un problème ouvert.

Dans cette thèse, le concept d'équilibre de satisfaction (satisfaction equlibrium, SE) a été formalisé et nous avons montré que cette notion d'équilibre est particulièrement adapté pour modéliser le problème de l'approvisionnement de qualité de service. Dans ce sens, nous avons présenté quelques applications. Cependant, des nombreux aspects théoriques restent à compléter. Par exemple, l'exploitation de la formulation de l'équilibre de satisfaction comme une inclusion de point fixe pour obtenir des résultats plus généraux sur l'existence ou l'unicité. Plus intéressant, une généralisation de la SE et le concept d'équilibre de satisfaction efficace (efficient satisfaction équilibre, ESE) à des jeux dynamiques, par exemple, des jeux stochastiques, reste à être formulée. Cette formulation dans les jeux dynamiques nous permettrait de modéliser la nature variante dans le temps de ce type de réseaux. D'un point de vue pratique, les algorithmes pour atteindre le SE et l'ESE d'une manière totalement décentralisée restent également à être conçus.

Contents

Abstract iii

Résumé iv

Présentation en français vi

1 Introduction 1
1.1 Technological Context and Challenges . . . 1
1.2 Contributions and Organization of the Thesis . . . 3
1.2.1 Contributions on Hierarchical Spectrum Access . . . 3
1.2.2 Contributions on Open Spectrum Access . . . 5
1.2.3 Contributions on the QoS Provisioning . . . 7
1.2.4 Other Contributions . . . 8
1.3 Notations . . . 9

2 Techniques for Hierarchical Spectrum Sharing 11
2.1 Opportunistic Interference Alignment Strategy . . . 12
2.1.1 System Model . . . 13
2.1.2 Interference Alignment Strategy . . . 16
2.1.3 Asymptotic Performance of the Secondary link . . . 22
2.1.4 Numerical Results . . . 26
2.2 Dynamic Bandwidth Limiting Strategy . . . 30
2.2.1 System Model . . . 32
2.2.2 Individual and Network Spectral Efficiency . . . 34
2.2.3 Asymptotic Network Spectral Efficiency . . . 35
2.2.4 Bandwidth Limiting Strategy . . . 36
2.2.5 Numerical Results . . . 37
2.3 Conclusions . . . 39

3 Equilibrium Analysis in Open Spectrum Sharing Games 41
3.1 Models . . . 43
3.1.1 Signal Model . . . 43
3.1.2 Game Model . . . 44
3.2 A Note on the Parallel Multiple Access Channel . . . 45
3.3 The Power Allocation Game $\mathcal{G}_{(a)}$ . . . 46
3.3.1 Existence of a Pure NE . . . 46

3.3.2 Uniqueness of the Pure NE 47
3.3.3 Dynamics Arising from the Best Response 48
3.4 The Channel Selection Game $\mathcal{G}_{(b)}$ 49
3.4.1 Existence of a Pure NE 49
3.4.2 Bounding the Number of Pure NE 50
3.4.3 Dynamics Arising from the Best Response 51
3.4.4 Dynamics Arising from the Fictitious Play 52
3.5 Analysis of Special Cases 54
3.5.1 The 2-Transmitter 2-Channel Case 54
3.5.2 A Braess Type Paradox 60
3.5.3 The Case of Large Systems 61
3.6 Numerical Examples 63
3.6.1 Impact of the SNR $\frac{p_{\max}}{\sigma_2}$ 63
3.6.2 Impact of Number of Transmitters per Dimension ($\frac{K}{S}$) 64
3.7 Conclusions 65

4 Learning Equilibria in Open Spectrum Sharing Games 66
4.1 System Model 68
4.2 Game Theoretic Model 68
4.2.1 Game Formulation 68
4.2.2 Nash Equilibrium and ϵ-Equilibrium 71
4.3 A Note on Stationary State-Independent Behavioral Strategies 71
4.3.1 Stationary State Independent Behavioral Strategies (SSI-BS) . 72
4.3.2 Achievable Performance with SSI-BS 72
4.3.3 Pertinence of the SSI-BS 73
4.4 Logit Equilibrium 74
4.4.1 Logit Equilibrium in SSI-BS 74
4.4.2 Existence in SSI-BS 76
4.4.3 Uniqueness in SSI-BS 76
4.5 Learning Logit Equilibria 77
4.5.1 Learning the Expected Performance 77
4.5.2 Learning the Equilibrium Probability Distributions 79
4.5.3 Joint Learning of the Expected Performance and ϵ-equilibrium Strategies 80
4.6 Convergence Analysis 82
4.6.1 Scenario 1: Homogeneous Learning Rates 83
4.6.2 Scenario 2: Heterogeneous Learning Rates 84
4.7 Applications 86
4.7.1 Logit Equilibrium and Interference Channels 86
4.7.2 Convergence in Finite Time 87
4.7.3 Impact of the Number of Choices 89
4.8 Conclusions 90

5 QoS Provisioning in Spectrum Sharing Games **91**
5.1 Problem Formulation . . . 93
5.2 Games in Satisfaction Form and Satisfaction Equilibrium . . . 94
5.2.1 Games in Satisfaction Form . . . 94
5.2.2 Extension in Mixed Strategies of the Satisfaction Form . . . 95
5.2.3 Epsilon-Satisfaction Equilibrium . . . 96
5.3 Existence and Uniqueness of the Satisfaction Equilibrium . . . 97
5.3.1 Existence of SE in Pure Strategies . . . 97
5.3.2 Existence of the SE in Mixed Strategies . . . 98
5.3.3 Existence of the Epsilon-Satisfaction Equilibrium . . . 99
5.3.4 Uniqueness of the SE . . . 100
5.4 Satisfaction Equilibrium and other Equilibrium Concepts . . . 100
5.4.1 Games in Normal Form and Satisfaction Form . . . 100
5.4.2 Satisfaction Equilibrium and Nash Equilibrium . . . 101
5.4.3 Satisfaction Equilibrium and Generalized Nash Equilibrium . . . 102
5.5 Equilibrium Selection and Efficient Satisfaction Equilibria . . . 103
5.5.1 Games in Efficient Satisfaction Form . . . 103
5.5.2 Existence of an ESE . . . 104
5.5.3 Uniqueness of the ESE . . . 106
5.6 Learning Satisfaction Equilibrium . . . 108
5.6.1 The Learning Problem Formulation . . . 108
5.6.2 Learning the SE in Pure Strategies . . . 109
5.6.3 Clipping Actions and SE . . . 111
5.7 Applications . . . 111
5.7.1 QoS Provisioning in the Interference Channel . . . 112
5.7.2 Clipping Actions in the Interference Channel . . . 113
5.7.3 Convergence Time to the SE . . . 114
5.8 Conclusions . . . 115

6 Conclusions and Perspectives **117**
6.1 Conclusions . . . 117
6.2 Perspectives . . . 120
6.2.1 On the Opportunistic Interference Alignment Strategy . . . 120
6.2.2 On the Bandwidth Limiting Strategy . . . 120
6.2.3 On the Equilibrium Analysis of Spectrum Sharing Games . . . 120
6.2.4 On the Schemes for Learning Equilibria . . . 121
6.2.5 On the Quality of Service Provisioning . . . 121

Appendix **133**

A Definitions **133**

B Proof of Lemma 2.1.4 **134**

C Proof of Proposition 2.1.1 **136**

D Proof of Prop. 3.5.1 141

E Proof of Prop. 3.5.2 144

F Proof of Prop. 3.5.5 145

G Proof of Prop. 3.5.6 146

H Proof of Lemma 4.3.3 149

I Proof of Lemma 4.5.1 152

J Proof of Theorem 4.5.2 158

K Proof of Lemma 4.5.3 159

List of Figures

1 Canal à interférence avec entrées et sorties multi dimensionelles (MIMO). xi

2 Débit du système opportuniste comme fonction du rapport signal sur bruit $\mathrm{RSB}_1 = \frac{p_{\max}}{\sigma_1^2}$. Le nombre d'antennes satisfait $N_t = N_r + 1$, avec $N_r \in \{3, 9\}$ et $RSB_1 = RSB_2$. La technique de formation de faisceaux (FF) à forçage de zero suit l'equation 5 avec puissance optimale. xv

3 Débit Asymptotique du système opportuniste avec AUP observé par simulation comme fonction du nombre d'antennes quand $N_r = N_t$ à différents niveaux de RSB. $\mathrm{RSB} = \frac{p_{\max}}{\sigma^2}$. Les lignes noires sont des asymptotes déterminées par le Th. 0.0.4. Les asymptotes du système primaire sont données par [30]. xvi

4 Équilibres de Nash comme fonction des rapports des canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$ pour le cas de deux joueurs et deux canaux dans les jeu $\mathcal{G}_{(a)}$. La fonction $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ est définie comme: $\psi(x) = 1 + \mathrm{SNR}\, x$. La correspondance de meilleure réponse $\mathrm{BR}_k(\mathbf{p}_{-k})$, pour tout $k \in \mathcal{K}$, est définie par (3.14). Dans cet exemple, nous avons choisi arbitrairement $\frac{\psi(g_{21})}{\psi(g_{12})} < \frac{\psi(g_{11})}{\psi(g_{22})}$. xxi

5 La fonction potentiel ϕ du jeu $\mathcal{G}_{(b)}$, avec $K = 2$ et $S = 2$. Le joueur 1 choisit les lignes et le joueur 2 choisit les colonnes. xxii

6 Équilibres de Nash comme fonction des rapports de canaux $\frac{g_{11}}{g_{12}}$ et $\frac{g_{21}}{g_{22}}$ pour le jeu à deux joueurs et deux canaux $\mathcal{G}_{(b)}$. La fonction $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ est définie comme suit: $\psi(x) = 1 + \mathrm{SNR}\, x$. Dans cet exemple, nous avons accordé arbitrairement $\psi(g_{11}) < \psi(g_{21})$. xxiii

2.1 Two-user MIMO interference channel. 14

2.2 Fraction of transmit opportunities in the RLNA (Def. 2.1.5), i.e., S_∞ (2.29) as function of the $\mathrm{SNR} = \frac{p_{1,\max}}{\sigma_1^2}$ and $\alpha_{11} = \frac{M_1}{N_1}$. Simulation results are obtained by using one realization of the matrix $\mathbf{H}_{11}$ when $N_1 = 10$. 27

2.3 Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\mathrm{SNR}_1 = \mathrm{SNR}_2$ when IA and ZFBF are implemented. The number of antennas satisfy $\alpha = \frac{N_t}{N_r} \approx \frac{5}{4}$, with $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r \in \{3, 9\}$ and $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_1^2}$, for all $i \in \{1, 2\}$. 29

2.4 Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\mathrm{SNR}_1 = \mathrm{SNR}_2$. The number of antennas satisfy $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r$, with $N_t = N_r$, and $N_r \in \{3, 6, 9\}$ and $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$, for all $i \in \{1, 2\}$. 29
2.5 Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$, with $i \in \{1, 2\}$. The number of antennas satisfy $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r$, with $N_r = N_t = 4$. 30
2.6 Asymptotic transmission rates per antenna of the opportunistic link as a function of the number of antennas when $N_r = N_t$ using uniform PA at different SNR levels $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$. Simulation results are obtained using one channel realization for matrices $\mathbf{H}_{ij}$ $\forall (i, j) \in \{1, 2\}^2$ and theoretical results using Prop. 2.1.1, ... 31
2.7 Vector multiple access channel with K transmitters and N available channels with $\mathbf{h_k} = (h_{k,1}, \ldots, h_{k,N})$ for all $k \in \mathcal{K}$. 33
2.8 Asymptotic transmission rate per channel (2.66) in bps/Hz as a function of the number of available channels N. Dashed lines are obtained from simulations considering $\frac{p_{\max}}{\sigma^2} = 20dB$ 37
2.9 Network Spectral Efficiency (2.65) in bps/Hz as a function of the maximum number of accessible channels L. Total number of available channels $N = 50$, and $\frac{p_{\max}}{\sigma^2} = 10dB$. 38
2.10 Optimal BL parameter L (2.80) as a function of the network load, $(\frac{K}{N})$. Total number of available channels $N = 50$, and $\frac{p_{\max}}{\sigma^2} = 10dB$. . 38
2.11 Network Spectral Efficiency (2.65) in bps/Hz as a function of the network load $(\frac{K}{N})$. Total number of available channels $N = 50$. 39

3.1 Nash equilibrium action profiles as a function of the channel ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$ for the two-player-two-channel game $\mathcal{G}_{(a)}$. The function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as follows: $\psi(x) = 1 + \mathrm{SNR}\, x$. The best response function $\mathrm{BR}_k(\mathbf{p}_{-k})$, for all $k \in \mathcal{K}$, is defined by (3.14). Here, it has been arbitrarly assumed that $\frac{\psi(g_{21})}{\psi(g_{12})} < \frac{\psi(g_{11})}{\psi(g_{22})}$. 56
3.2 Potential function ϕ of the game $\mathcal{G}_{(b)}$, with $K = 2$ and $S = 2$. Player 1 chooses rows and player 2 chooses columns. 57
3.3 Nash equilibrium action profiles as a function of the ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$ of the channel realizations for the two-player-two-channel game $\mathcal{G}_{(b)}$. The function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as follows: $\psi(x) = 1 + \mathrm{SNR}\, x$. Here, it has been arbitrarily assumed that $\psi(g_{11}) < \psi(g_{21})$. 58
3.4 (a) Network spectral efficiency as a function of the system load $\eta = \frac{K}{S}$ for different $\mathrm{SNR} = \frac{p_{\max}}{\sigma^2}$ levels in dBs. (b) Network spectral efficiency as a function of the $\mathrm{SNR} = \frac{p_{\max}}{\sigma^2}$ in dBs. for the case of $\eta = \frac{K}{S} \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, with $K = 10$. 63
3.5 Probability of observing a specific number of NE in the game $\mathcal{G}_{(b)}$. . . 64

3.6 Fraction of players transmitting over channel s, with $s \in \mathcal{S}$, calculated using Monte-Carlo simulations and using Eq. (3.47) for a network with $S = 6$ channels, $\eta = 10$ players per channel, with $\mathbf{b} = (b_s)_{\forall s \in \mathcal{S}} = (0.25, 0.11, 0.20, 0.05, 0.25, 0.14)$, and SNR $= \frac{p_{\max}}{N_0 B} = 10$ dB. 65

4.1 Average sum spectral efficiency in the two-transmitter two-receiver two-channel game, when transmitters are limited to channel selection. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{(\frac{3}{4})}}$, $\alpha_2(n) = \frac{1}{n^{(\frac{2}{3})}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs. 88

4.2 Average sum spectral efficiency in the two-transmitter two-receiver two-channel game, when transmitters are limited to channel selection. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{(\frac{3}{4})}}$, $\alpha_2(n) = \frac{1}{n^{(\frac{2}{3})}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs. 89

4.3 Achieved sum spectral efficiency in the two-transmitter two-receiver N-channel game, when transmitters are limited to use only one channel or any combination of adjacent channels. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{(\frac{3}{4})}}$, $\alpha_2(n) = \frac{1}{n^{(\frac{2}{3})}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs. 90

5.1 Game in normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$, with $\mathcal{K} = \{1, 2\}$, $\mathcal{A}_k = \{A_k^{(1)}, A_k^{(2)}\}$, for all $k \in \mathcal{K}$. Player 1 chooses rows and player 2 chooses columns. In a pair $(v_1, v_2) \in \{0, 1\}^2$, v_1 and v_2 are the utilities obtained by player 1 and 2, respectively. 101

5.2 Achievable (Shannon) transmission rates $(u_1(p_1, p_2), u_2(p_1, p_2))$, for all $(p_1, p_2) \in \mathcal{A}_1 \times \mathcal{A}_2$, with SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels. 113

5.3 Histogram of the event of convergence or non-convergence of the learning algorithm (Alg. 1) in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$. Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels. 114

5.4 Histogram of the convergence time to a SE in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ using the algorithm (Alg. 1). Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels. 115

5.5 Instantaneous achieved rates of transmitter 1 (red) and 2 (blue).Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels. 116

Chapter 1

Introduction

In this chapter, we present a general description of the problem of spectrum sharing in decentralized self-configuring networks. In particular, we identify the main technological challenges and we describe the main contributions presented in this thesis. Finally, we conclude by describing the notation used in the sequel.

1.1 Technological Context and Challenges

A decentralized self-configuring network (DSCN) is basically an infrastructure-less network where radio devices autonomously determine their own transmit/receive configuration in order to guarantee a reliable communication. For this purpose, radio devices are often equipped with spectrum sensing and self-configuring capabilities. Therefore, this kind of radio devices, often called cognitive radios (CR), are able to identify the unused radio resources and tune their transmit/receive configuration to efficiently exploit such resources. For instance, transmission parameters such as power allocation policies, modulation-coding schemes and antenna array configurations can be dynamically tuned by each radio device to adapt to its current environment.

The underlying feature of DCSNs is that transmitters communicate with their respective receivers without the control of a central authority, for instance, a base station. Thus, the main limitation of these networks is the mutual interference arising from the uncoordinated interaction of a set of transmitters exploiting a common set of radio resources. Indeed, this is the reason why the analysis and design of spectrum sharing techniques plays a central role in this scenario. Here, among all the constraints for achieving an optimal exploitation of the spectrum, we mention two facts. First, radio devices must determine their own transmit/receive configuration relying only on local information. Second, the global network topology is constantly changing. Additionally, these networks must be quickly deployable, as well as robust to attack and failure in critic application environments. These last requirements arise from the fact that DSCN are typically used in many military, law enforcement, disaster relief, space, and indoor/outdoor commercial applications.

In general, spectrum sharing in the context of DSCN might follow two different sets

of rules depending on whether they operate over licensed or unlicensed bands [122]. In the case of licensed bands, we say that the DCSN follows a hierarchical spectrum access (HSA) and in the case of non-licensed bands, we say it follows an open spectrum access (OSA).
In HSA, DCSNs operate under the condition that the additional interference it generates over the pre-existing systems is either inexistent (spectrum overlay) or below a specific threshold (spectrum underlay) [17]. A typical example of spectrum overlay is precisely the idea behind of CR, where only spectral resources left over by pre-existing systems (spectrum access opportunities, SAO) are used, and thus, no additional interference is generated. Depending on the multiple access technique of the primary system, a SAO consists of non-occupied time slots in time division multiple access (TDMA), frequency bands in frequency division multiple access (FDMA), spatial directions in spatial division multiple access (SDMA), tones in orthogonal frequency division multiple access (OFDMA), spreading codes in code division multiple access (CDMA) or a combination of any of those. In the context of spectrum underlay, a classical example is the ultra-wide band (UWB) modulation. Here, the interference produced by the radio devices using UWB modulation does not represent a significant additional interference to the legacy system, and thus, such coexistence is tolerated.
In OSA, each radio device has the same rights to access the spectrum at any time. This is particularly the case of unlicensed bands (e.g., the industrial, scientific and medical (ISM) band [2.400, 2.500] GHz). Radio devices operating in unlicensed bands include cordless telephones, wireless sensors and devices operating under the standards of Wi-Fi (IEEE 802.11), Zig-Bee (IEEE 802.15.4), and Bluetooth (IEEE 802.15.1).
In this thesis, we focus on the case of both HSA and OSA. In particular, we highlight the fact that in HSA, once either the available SAOs are reliably identified (spectrum overlay) or the instantaneous interference level produced over the primary system is known by all the radio devices in the DCSN (spectrum underlay), the analysis of the spectrum sharing is identical in both HSA and OSA. The important remark here is that, two or more radio devices in the DCSN exploiting the same radio resources are subject to mutual interference, independently of whether they operate under HSA or OSA policies. The main difference between HSA and OSA is that in the former, the radio resources are available only the period the primary system does not use it, while in the latter, radio resources are always available. However, in this thesis, we do not consider this constraint, and we assume that available radio resources, in HSA, remain available a period longer than the communication duration in the DSCN. Under this assumption, the common problem with either HSA (overlay and underlay spectrum) and OSA is that there is a group of terminals simultaneously exploiting the same radio resources and thus, subject to mutual interference. In the following, we use the generic term spectrum sharing to refer to both HSA and OSA under the conditions stated above.
Within this framework, the spectrum access problem, as treated in this thesis, can be summarized by the following set of questions:

(*i*) What is the optimal individual performance radio devices can observe in DSCN? (*ii*) What is the optimal behavior a radio device must adopt to achieve an optimal individual performance given the environment and actual configuration of all the other devices?
To tackle the first question, game theory, a branch of mathematics which studies the interaction between several decision-makers, is the dominant paradigm followed in this thesis [46,50,51,79,106]. In particular, we use the idea of equilibrium [46], which basically determines the stable states of a given DSCN. Here, stability is interpreted as a network state where the transmit/receive configuration of each radio device is optimal with respect to all the other radio devices. Thus, none of the radio devices improves its performance by unilaterally deviating from the equilibrium state. Regarding question (*ii*), by the term behavior, we refer to the policy a radio device uses to select the transmit configuration based on its available information about the network. Here, an important remark is that some of the network equilibria can be achieved as a result of an iterative interacting process, similar to a learning process. Thus, as an additional tool to deal with question (*ii*), we use some elements from machine learning theory.
In the following, we describe the main contributions made in this thesis.

1.2 Contributions and Organization of the Thesis

The contributions made in this thesis fall into three main areas: (*i*) performance analysis and design of techniques for hierarchical spectrum access (HSA), (*ii*) performance analysis and design of techniques for open spectrum access (OSA), and (*iii*) mechanisms for quality of service (QoS) provisioning in both HSA and OSA networks. The contributions in HSA are presented in Chapter 2. The contributions in OSA are presented in both Chapter 3 and Chapter 4. The contributions regarding the QoS provisioning are presented in Chapter 5. In the following, we provide a brief description of each of these contributions.

1.2.1 Contributions on Hierarchical Spectrum Access

In the context of HSA, two contributions are presented. The first one, in Sec. 2.1, concerns an opportunistic scheme for inference alignment in MIMO cognitive networks [77,78,87]. The second contribution, in Sec. 2.2, consists in a technique for improving the spectral efficiency of cognitive networks by strategically modifying the number of channels cognitive radios are let to use [75]. In the following, we describe in detail each contribution.

An Opportunisitic Inteference Alignment Strategy

The opportunistic interference alignment (OIA) technique presented in this thesis allows an opportunistic multiple input multiple output (MIMO) link (secondary) to harmlessly coexist with another MIMO link (primary) in the same frequency band.

This contribution is presented in the context of the interference channel. However, as recently shown in [45], it easily extends to other network topologies. The main idea behind this novel technique is the following. Assuming perfect channel knowledge at the primary receiver and transmitter, capacity is achieved by transmitting along the spatial directions (SD) associated with the singular values of its channel matrix using a water-filling power allocation (PA) scheme. Often, power limitations lead the primary transmitter to leave some of its SD unused. Here, it is shown that the opportunistic link can transmit its own data if it is possible to align the interference produced on the primary link with such unused SDs. Both the processing scheme to perform IA and the corresponding PA scheme which maximizes the transmission rate of the opportunistic link are presented. The asymptotes of the achievable transmission rates of the opportunistic link are obtained in the regime of large numbers of antennas. Using this result, it is proved that depending on the signal-to-noise ratio and the number of transmit and receive antennas of the primary and opportunistic links, both systems can achieve transmission rates of the same order. The publications concerning this contribution are listed hereunder:

- Perlaza, S. M. and Fawaz, N. and Lasaulce, S. and Debbah, M., "From Spectrum Pooling to Space Pooling: Opportunistic Interference Alignment in MIMO Cognitive Networks". IEEE Trans. in Signal Processing, vol.58, no.7, pp.3728-3741, July 2010.

- Perlaza, S. M., and Debbah, M. and Lasaulce, S. and Chaufray, J-M., "Opportunistic interference alignment in MIMO interference channels", in IEEE Intl. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, September, 2008. **(invited paper)**.

- Perlaza, S. M. and Fawaz, N. and Lasaulce, S. and Debbah, M., "Alignement d'interférence opportuniste avec des terminaux multi-antennes", in proc. of the Gretsi Conference. Dijon, France, 2009.

A Bandwidth Limiting Strategy

The second contribution of this thesis, in the context of HSA, is a technique that consists in strategically limiting the number of frequency bands that transmitters can use, in order to increase the network spectral efficiency. This contribution is presented in the context of parallel multiple channels, however, it easily extends to other network topologies. This technique, to which we refer as bandwidth limiting (BL), applies to networks where the priority of each transmitter depends on its order of arrival. For instance, the first arrived radio device can use the whole available bandwidth, the second uses what is left by the first one, and so on. Assuming that each transmitter maximizes its own data rate by water-filling over the available frequency bands, we show the existence of an optimal number of dimensions that a transmitter must use in order to maximize the network performance measured in terms of spectral efficiency. We provide a closed form expression for the optimal number of accessible bands. Such an optimum point, depends on the number of

active transmitters, the number of available frequency bands and the different signal-to-noise ratios. The impact of each of this parameters over the network spectral efficiency is studied using numerical examples.
The publications concerning this contribution are listed hereunder:

- Perlaza, S. M., and Debbah, M. and Lasaulce, S. and Bogucka, H. "On the benefits of bandwidth limiting in vector multiple access channels", in EURASIP/IEEE Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Communications(CROWNCOM), Hannover, Germany, June, 2009. **(Best Student Paper Award)**

- Corvino, V. and Moretti, M. and Perlaza, S. M. and Debbah, M. and Lasaulce, S. and Jouini, W. and Palicot, J. and Moy, C. and Serrador, A. and Bogucka, H. and Sroka, P. and Rodrigues, E. B. and López-Benítez, M. and Umbert, A. and Casadevall, F. and Pérez-Romero, J. "Definition and evaluation of Joint Radio Resource Management and Advanced Spectrum Management Algorithms", NEWCOM++ Network of Excellence, Deliverable DR9.2, January 12, 2010.

1.2.2 Contributions on Open Spectrum Access

In the context of OSA, two contributions are presented in this thesis. First, in Chapter 3, the (Nash) equilibrium analysis of a decentralized parallel multiple access channel is presented [73, 80, 86]. This scenario models the case where several transmitters aim to communicate with a single receiver at the maximum spectral efficiency sharing a common set of available frequency bands. The second contribution, in Chapter 4, consists in a methodology for designing behavioral rules or learning dynamics that allow radio devices to achieve equilibrium as a result of an iterative interaction similar to a learning process [81, 82, 84]. These contributions are detailed here under.

Nash Equilibrium in Open Spectrum Access Games

The first contribution in the context of OSA consists of the analysis of two spectrum sharing problems, namely the power allocation (PA) and channel selection (CS) problems. The former considers the spectrum sharing case where transmitters can simultaneously use several channels, while the latter considers the case where transmitters use only one channel at a time. These problems are studied in the context of parallel multiple access channels, where transmitters selfishly maximize their individual spectral efficiency. In the context of parallel MAC, both problems are modeled by non-cooperative games and a thorough Nash equilibrium (NE) analysis is conducted. The corresponding games have some attractive properties not available for other channels like multiple input multiple output interference channels. For instance, the studied games are potential games and thus, the existence of at least one NE is guaranteed. Moreover, a unique NE is observed almost surely in the PA game. In the CS game, by using a graph-theoretic interpretation of the

set of NE, it is shown that multiple pure NE exist and their multiplicity is studied. Convergence of certain important dynamics like the best-response dynamics and fictitious-play are discussed. Comparing the performance at the NE of both games in a special case, the existence of a Braess-type paradox is proved: having more choices in terms of power allocation policies for the transmitters can lead to a worse outcome in terms of network spectral efficiency. The results presented here are network topology dependent. A similar analysis has been carried out for the case of parallel interference channels [90,91], however, these results are not included in this thesis.
The publications concerning this contribution are listed hereunder:

- Perlaza, S. M. and Florez, V. Q. and Tembine, H. and Lasaulce, S., "On the Convergence of Fictitious Play in Channel Selection Games", IEEE Latin America Transactions. April 2011. **(Invited paper)**

- Perlaza, S. M. and Tembine, H. and Lasaulce, S. and Florez, V. Q., "On the Fictitious Play and Channel Selection Games", in proc. of the Latin-American Conference on Communications (LATINCOM), Bogotá, Colombia, September, 2010.

- Perlaza, S. M., and Belmega, E. V. and Lasaulce, S. and Debbah, M., "On the base station selection and base station sharing in self-configuring networks", in Fourth International Conference on Performance Evaluation Methodologies and Tools, Pisa, italy, October, 2009. **(Invited paper)**.

Learning Techniques for Achieving NE in Open Spectrum Sharing Games

The second contribution in the context of OSA consists in a general framework for designing behavioral rules for radio devices aiming to achieve individually optimal performance. The underlying assumptions of this framework are the following: (i) the state of the network is time-varying and it is modeled by a set of random variables; (ii) radio devices are interested in their long-term average performance rather than instantaneous performance; (iii) the information that each radio device possesses about the network at a given instant is a numerical observation of its own performance. Following these assumptions, we present a generic stochastic game which models most of radio resource sharing scenarios in interference limited systems. Our main contribution consists of a family of behavioral rules that allow radio devices to simultaneously perform two tasks. First, to build an estimation of their own individual average performance associated with each of their actions. Second, to achieve an epsilon-Nash equilibrium of the corresponding stochastic game by using such estimations. The additional performance estimation helps to solve the problem encountered in behavioral rules based on reinforcement learning, where convergence is observed but the final network configuration does not correspond to an NE. Following the proposed rules, if convergence is observed, the final configuration corresponds to a logit equilibrium. A thorough analysis of the convergence

properties of these behavioral rules is presented. In particular, a simple parallel interference channel scenario is used to compare the proposed algorithm with existing learning dynamics such as best response dynamics, fictitious play dynamics, regret matching learning and reinforcement learning.
The publications concerning this contribution are listed hereunder:

- Rose, L. and Perlaza, S. M. and Lasaulce, S. and Debbah, M., "Learning Equilibria with Partial Information in Wireless Networks", Submitted to the IEEE Communications Magazine, Special Issue in Game Theory for Wireless Communications, August, 2011.

- Perlaza, S. M. and Lasaulce, S. and Tembine, H. and Debbah, M., "Learning to Use the Spectrum in Self-Configuring Heterogeneous Networks: A Logit Equilibrium Approach", in proc. of the 4th International ICST Workshop on Game Theory in Communication Networks (GAMECOMM), Paris, France, May, 2011.

- Perlaza, S. M. and Tembine, H. and Lasaulce, S., "How can Ignorant but Patient Cognitive Terminals Learn Their Strategy and Utility?", in proc. of the IEEE Intl. Workshop on Signal Processing Advances for Wireless Communications (SPAWC), Marrakesh, Morroco, June, 2010.

1.2.3 Contributions on the QoS Provisioning

In the context of QoS provisioning, the main contribution is presented in Chapter 5 and it consists in the formalization of a particular equilibrium concept, namely the satisfaction equilibrium (SE). In contrast to existing equilibrium notions, for instance Nash equilibrium (NE) and generalized NE (GNE), in the SE, the idea of performance optimization in the sense of utility maximization or cost minimization does not exist. The concept of SE relies on the fact that players might be either satisfied or unsatisfied with their achieved performance. At the SE, if it exists, all players are satisfied. This notion of equilibrium perfectly models the problem of QoS provisioning in decentralized self-configuring networks. Here, radio devices are satisfied if they are able to provide the requested QoS. Within this framework, the concept of SE is formalized for both pure and mixed strategies. In both cases, sufficient conditions for the existence and uniqueness of the SE are presented. When multiple SE exist, we introduce the idea of effort or cost of satisfaction and we propose a refinement of the SE, namely the efficient SE (ESE). At the ESE, all players adopt the action which requires the lowest effort for satisfaction. A learning method that allows radio devices to achieve a SE in pure strategies in finite time is also presented. In contrast to existing methods for achieving (Nash) equilibria, this method only requires one bit feedback from the receiver at every learning stage. Finally, the advantages of the SE with respect to NE and GNE are highlighted using a simple power control game in the interference channel.
The publications concerning this contribution are listed hereunder:

- Perlaza, S. M. and Tembine, H. and Lasaulce, S. and Debbah, M., "A General Framework for Quality-Of-Service Provisioning in Decentralized Networks". Submitted to the IEEE Journal in Selected Topics in Signal Processing. Special Issue in Game Theory for Signal Processing. 2011

- Perlaza, S. M. and Tembine, H. and Lasaulce, S. and Debbah, M., "QoS Provisioning in Self-Configuring Wireless Networks: Beyond Nash Equilibrium". IEEE Communications Society R-Letters, vol. 2, no. 1, February 2011. **(Invited Paper)**

- Perlaza, S. M. and Tembine, H. and Lasaulce, S. and Debbah, M., "Satisfaction Equilibrium: A General Framework for QoS Provisioning in Self-Configuring Networks", in proc. of the IEEE Global Communications Conference (GLOBECOM),Miami, USA, December, 2010. **(Candidate for Best Conference Paper Award of the IEEE MMTC Review Board Chair in 2011)**

1.2.4 Other Contributions

Other contributions consist of applications of the theoretical frameworks presented in this thesis. Such contributions have been published in cooperation with other authors in the context of collaborations or short-term visits to other laboratories. Other contributions have been kept in the form of patents and they are held by France Telecom. Hereunder, we present the list of publications which are not included in this thesis.

- Bennis, M. and Perlaza, S. M.,"Decentralized Cross-Tier Interference Mitigation in Cognitive Femtocell Networks", in proc. of the IEEE International Conference on Communications (ICC), Kioto, Japan, June, 2011.

- Rose, L. and Perlaza, S. M., and Debbah, M., "On the Nash Equilibria in Decentralized Parallel Interference Channels", in proc. of the IEEE ICC 2011 Workshop on Game Theory and Resource Allocation for 4G, Kioto, Japan, June, 2011.

- Bennis, M. and Perlaza, S.M. and Debbah, M. "Game Theory and Femtocell Communications: Making Network Deployment Feasible" in Saeed, R. A. and Chaudhari, B. S. (Editors), "Femtocell Communications: Business Opportunities and Deployment Challenges". IGI Global, USA, 2011.

- Perlaza, S. M. and Lasaulce, S. and Debbah, M. and Chaufray, J-M., "Game Theory for Dynamic Spectrum Access", in Y. Zhang, J. Zheng, and H.-H. Chen (Eds.), Cognitive Radio Networks: Architectures, Protocols and Standards, Auerbach Publications, USA, 2010.

- Perlaza, S. M., and Cottatellucci, L. and Debbah, M., "A Game Theoretic Framework for Decentralized Power Allocation in IDMA Systems", in IEEE

Intl. Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Cannes, France, September, 2008.

Other contributions in the form of patents are listed hereunder.

- Orange Labs - France Telecom, (Perlaza, S. M and G. Salingue). “Procédé de sélection de canal par un émetteur, procédé et dispositif d'émission de données et programme d'ordinateur associés”, France, 2010.

- Orange Labs - France Telecom, (Perlaza, S. M and S. Lasaulce and G. Salingue). “Communications between self-configuring radios using dynamic updates of the transmission configuration”, France, 2010

1.3 Notations

In the sequel of this thesis, matrices and vectors are respectively denoted by boldface upper case symbols and boldface lower case symbols. An $N \times K$ matrix with ones on its main diagonal and zeros on its off-diagonal entries is denoted by $\mathbf{I}_{N\times K}$, while the identity matrix of size N is simply denoted by $\mathbf{I}_N$. An $N \times K$ matrix with zeros in all its entries (null matrix) is denoted by $\mathbf{0}_{N\times K}$. Matrices $\mathbf{X}^T$ and $\mathbf{X}^H$ are the transpose and Hermitian transpose of matrix $\mathbf{X}$, respectively. The determinant of matrix $\mathbf{X}$ is denoted by $|\mathbf{X}|$. The expectation operator is denoted by $\mathbb{E}[.]$. The indicator function associated with a given set $\mathcal{A}$ is denoted by $\mathbb{1}_{\mathcal{A}}(.)$, and defined by $\mathbb{1}_{\mathcal{A}}(x) = 1$ (resp. 0) if $x \in \mathcal{A}$ (resp. $x \notin \mathcal{A}$). The indicator function associated to a condition is denoted by $\mathbb{1}_{\{condition\}}$ and it equals 1 (resp. 0) when *condition* is true (resp. false). The Heaviside step function and the Dirac delta function are respectively denoted by $\mu(\cdot)$ and $\delta(\cdot)$. The symbols $\mathbb{N}$, $\mathbb{R}$, and $\mathbb{C}$ denote the sets of non-negative integers, real numbers, and complex numbers, respectively. The subsets $[0, +\infty[$ and $]-\infty, 0]$ are denoted by $\mathbb{R}^+$ and $\mathbb{R}^-$, respectively. The operator $(x)^+$ with $x \in \mathbb{R}$ is equivalent to the operation $\max(0, x)$. Let $\mathbf{A}$ be an n-dimensional square matrix with real eigenvalues $\lambda_{A,1}, \ldots, \lambda_{A,n}$. We define the empirical eigenvalue distribution of $\mathbf{A}$ by $F_A^{(n)}(\cdot) \triangleq \frac{1}{n}\sum_{i=1}^{n} \mu(\lambda - \lambda_{A,i})$, and, when it exists, we denote $f_A^{(n)}(\lambda)$ the associated eigenvalue probability density function, where $F_{\mathbf{A}}(\cdot)$ and $f_{\mathbf{A}}(\cdot)$ are respectively the associated limiting eigenvalue distribution and probability density function when $n \to +\infty$.

The sets are denoted by calligraphic letters. Let $\mathcal{X}$ be a finite set. We denote by $\triangle(\mathcal{X})$ the unit simplex over the elements of $\mathcal{X}$, that is, the set of all probability distributions over the elements of the set $\mathcal{X}$. The cardinality of $\mathcal{X}$ is denoted by $|\mathcal{X}| \in \mathbb{N}$ and the set of all subsets of $\mathcal{X}$ including the set $\mathcal{X}$ itself is denoted by $2^{\mathcal{X}}$. We denote by $\{\mathbf{e}_1^{(N)}, \ldots \mathbf{e}_N^{(N)}\}$ the set of vectors of the canonical base spanning the space of the N-dimensional real vectors. Here, $\forall n \in \{1, \ldots, N\}$, $\mathbf{e}_n^{(N)} = \left(e_{n,1}^{(N)}, \ldots, e_{n,N}^{(N)}\right)$, and $\forall s \in \{1, \ldots, N\} \setminus \{n\}$, $e_{n,s}^{(N)} = 0$ and $e_{n,n}^{(N)} = 1$. Given a vector $\mathbf{a} = (a_1, \ldots, a_N)$ in a given space of dimension $N \in \mathbb{N}$, we denoted by $\mathbf{a}_{-n}$, with $n \in \{1, \ldots, N\}$, the vector $\mathbf{a}_{-n} = (a_1, \ldots, a_{n-1}, a_{n+1}, \ldots, a_N)$ in the corresponding space of dimension

$N-1$. With a slight abuse of notation, we often write $\mathbf{a} = (a_n, \mathbf{a}_{-n})$ to highlight the n-th component of the vector $\mathbf{a}$.

Chapter 2

Techniques for Hierarchical Spectrum Sharing

In this chapter, we study the performance of two different HSA schemes. In the first case, we consider a two-level hierarchy and in the second one, we consider a multi-level hierarchy. In both cases, the underlying assumptions of our analysis are the following: (i) a radio device coexists in the same spectrum block with a higher priority radio device, only if the additional interference it generates is zero, (ii) all radio devices have different priorities. The first scenario is studied in the context of the multiple input multiple output (MIMO) cognitive interference channel and the second one is studied in the context of a parallel multiple access channel. However, the same analysis holds for other network topologies. Here, the assumption of a particular scenario is done in order to facilitate the presentation and the interpretation of the obtained results.

In the case of the two-level hierarchy scenario, a non-cooperative interference alignment (IA) technique. This novel IA technique allows an opportunistic MIMO link (low priority) to harmlessly coexist with another MIMO link (top priority) in the same frequency band. The asymptotes of the achievable transmission rates of the opportunistic link are obtained in the regime of large numbers of antennas. Using this result, it is proved that depending on the signal-to-noise ratio (SNR) and the number of transmit and receive antennas of both links, non-negligible transmission rates can be obtained by the opportunistic link.

In the case of the multi-level hierarchy, it is shown that the network spectral efficiency can be improved by limiting the number of channels each transmitter is allowed to use. We refer to this technique as bandwidth limiting (BL). In the context of a parallel multiple access channel, the asymptotes of the individual and network spectral efficiency are obtained in the regime of large number of channels. Later, based on such asymptotical analysis, the optimal number of channels each transmitter must access to maximize the network spectral efficiency is determined. In this case, such an optimum operating point depends mainly on the network load (transmitters per channel) and the different signal to noise ratios of the different communications. Our numerical analysis shows that, the asymptotic results hold even for a finite number of radio devices and available channels.

2.1 Opportunistic Interference Alignment Strategy

The concept of HSA has been described in Sec. 1.1. The main idea is to let a class of radio devices, called secondary systems, opportunistically access certain portions of spectrum left unused by the primary system, at a given time or geographical area [36]. These pieces of unused spectrum, known as white-spaces, appear mainly when either transmissions in the primary network are sporadic, i.e., there are periods over which no transmission takes place, or there is no network infrastructure for the primary system in a given area, for instance, when there is no primary network coverage in a certain region. In the case of dense networks, observing a white-space might be a rare and short-lasting event. In fact, the idea of cognitive radio as presented in [36] (i.e., spectrum pooling), depends on the existence of such white-spaces [33]. In the absence of those spectrum holes, secondary systems are unable to transmit without producing additional interference on the primary systems. One solution to this situation has been provided recently under the name of interference alignment (IA). Basically, IA refers to the construction of signals such that the resulting interference signal lies in a subspace orthogonal to the one spanned by the signal of interest at each receiver. The IA concept was independently introduced by several authors [20, 52, 53, 116]. Recently, IA has become an important tool to study several aspects of the interference channel, namely its degrees of freedom [19, 20, 38]. The feasibility and implementation issues of IA regarding mainly the required channel state information (CSI) has been also extensively studied [32, 88, 110, 111].
In this section, we study an IA scheme named opportunistic IA (OIA) [76]. The idea behind OIA can be briefly described as follows. The primary link is modeled by a single-user MIMO channel since it must operate free of any additional interference produced by secondary systems. Then, assuming perfect CSI at both transmitter and receiver ends, capacity is achieved by implementing a water-filling power allocation (PA) scheme [109] over the spatial directions associated with the singular values of its channel transfer matrix. Interestingly, even if the primary transmitters maximize their transmission rates, power limitations generally lead them to leave some of their spatial directions (SD) unused. The unused SD can therefore be reused by another system operating in the same frequency band. Indeed, an opportunistic transmitter can send its own data to its respective receiver by processing its signal in such a way that the interference produced on the primary link impairs only the unused SDs. Hence, these spatial resources can be very useful for a secondary system when the available spectral resources are fully exploited over a certain period in a geographical area. The idea of OIA, as described above, was first introduced in [76] considering a very restrictive scenario, e.g., both primary and secondary devices have the same number of antennas and same power budget. Here, we consider a more general framework where devices have different number of antennas, different power budgets and no conditions are impossed over the channel transfer matrices (In [76], full rank condition was impossed over certain matrices).
The rest of this section is structured as follows. First, the system model, which

consists of an interference channel with MIMO links, is introduced in Sec. 2.1.1. Then, our aim in Sec. 2.1.2 is twofold. First, an analysis of the feasibility of the OIA scheme is provided. For this purpose, the existence of transmit opportunities (SD left unused by the primary system) is studied. The average number of transmit opportunities is expressed as a function of the number of antennas at both the primary and secondary terminals. Second, the proposed interference alignment technique and power allocation (PA) policy at the secondary transmitter are described. In Sec. 2.1.3, tools from random matrix theory for large systems are used to analyze the achievable transmission rate of the opportunistic transmitter when no optimization is performed over its input covariance matrix. We illustrate our theoretical results by simulations in Sec. 2.1.4. Therein, it is shown that our approach allows the secondary link to achieve transmission rates of the same order as those of the primary link. Finally, in Sec. 2.3 we state our conclusions and provide possible extensions of this work.

2.1.1 System Model

We consider two unidirectional links simultaneously operating in the same frequency band and producing mutual interference as shown in Fig. 2.1. The first transmitter-receiver pair $(\mathrm{Tx}_1, \mathrm{Rx}_1)$ is the primary link. The pair $(\mathrm{Tx}_2, \mathrm{Rx}_2)$ is an opportunistic link subject to the strict constraint that the primary link must transmit at a rate equivalent to its single-user capacity. Denote by N_i and M_i, with $i = 1$ (resp. $i = 2$), the number of antennas at the primary (resp. secondary) receiver and transmitter, respectively. Each transmitter sends independent messages only to its respective receiver and no cooperation between them is allowed, i.e., there is no message exchange between transmitters. This scenario is known as the MIMO interference channel (IC) [102, 114] with private messages. A private message is a message from a given source to a given destination: only one destination node is able to decode it. Indeed, we do not consider the case of common messages which would be generated by a given source in order to be decoded by several destination nodes.

Here, we assume the channel transfer matrices between different nodes to be fixed over the whole duration of the transmission. The channel transfer matrix from transmitter $j \in \{1, 2\}$ to receiver $i \in \{1, 2\}$ is an $N_i \times M_j$ matrix denoted by $\mathbf{H}_{ij}$ which corresponds to the realization of a random matrix with independent and identically distributed (i.i.d.) complex Gaussian circularly symmetric entries with zero mean and variance $\frac{1}{M_j}$, which implies

$$\forall (i, j) \in \{1, 2\}^2, \quad \mathrm{Trace}\left(\mathbb{E}\left[\mathbf{H}_{ij}\, \mathbf{H}_{ij}^H\right]\right) = N_i. \tag{2.1}$$

The L_i symbols transmitter i is able to simultaneously transmit, denoted by $s_{i,1}, \ldots, s_{i,L_i}$, are represented by the vector $\mathbf{s}_i = (s_{i,1}, \ldots, s_{i,L_i})^T$. We assume that $\forall i \in \{1, 2\}$ symbols $s_{i,1}, \ldots, s_{i,L_i}$ are i.i.d. zero-mean circularly-symmetric complex Gaussian variables. In our model, transmitter i processes its symbols using a matrix $\mathbf{V}_i$ to construct its transmitted signal $\mathbf{V}_i \mathbf{s}_i$. Therefore, the matrix $\mathbf{V}_i$ is called pre-processing

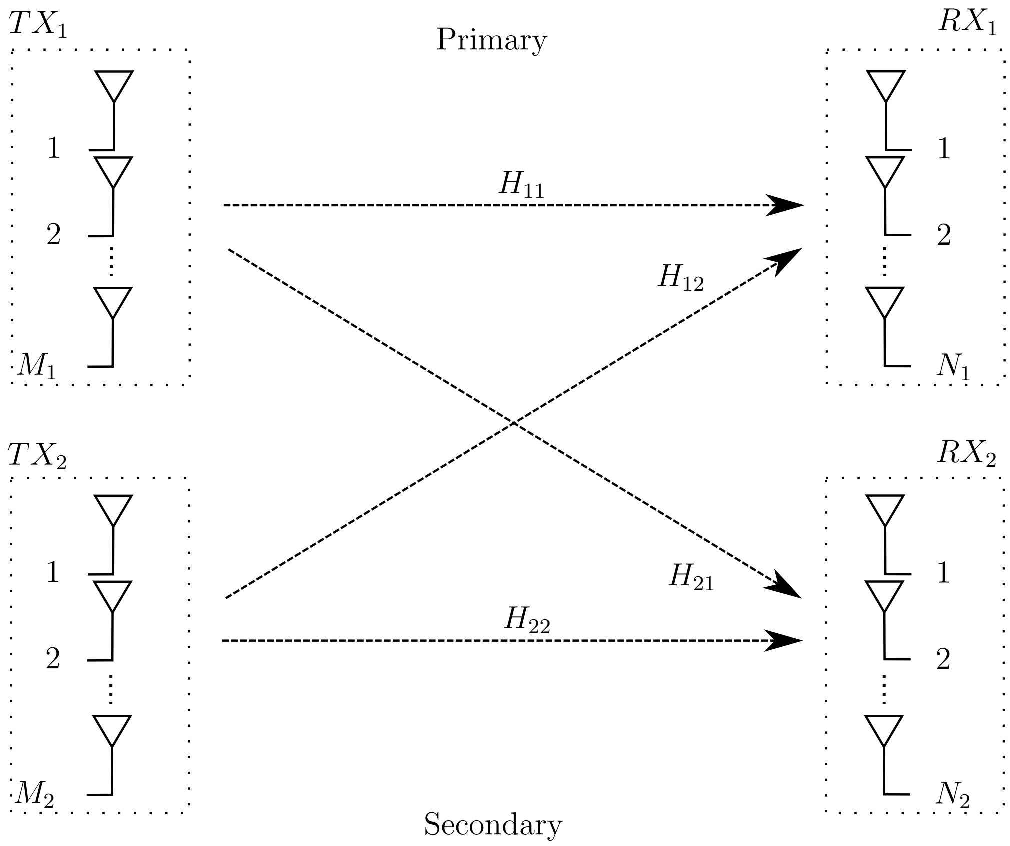

Figure 2.1: Two-user MIMO interference channel.

matrix. Following a matrix notation, the primary and secondary received signals, represented by the $N_i \times 1$ column-vectors $\mathbf{r}_i$, with $i \in \{1,2\}$, can be written as

$$\begin{pmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{V}_1\mathbf{s}_1 \\ \mathbf{V}_2\mathbf{s}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{pmatrix}, \tag{2.2}$$

where $\mathbf{n}_i$ is an N_i-dimensional vector representing noise effects at receiver i with entries modeled by an additive white Gaussian noise (AWGN) process with zero mean and variance σ_i^2, i.e.,$\forall i \in \{1,2\}$, $\mathbb{E}\left[\mathbf{n}_i\mathbf{n}_i^H\right] = \sigma_i^2\mathbf{I}_{N_i}$. At transmitter $i \in \{1,2\}$, the $L_i \times L_i$ power allocation matrix $\mathbf{P}_i$ is defined by the input covariance matrix $\mathbf{P}_i = \mathbb{E}\left[\mathbf{s}_i\mathbf{s}_i^H\right]$. Note that symbols $s_{i,1} \ldots, s_{i,L_i}$, $\forall i \in \{1,2\}$ are mutually independent and zero-mean, thus, the PA matrices can be written as diagonal matrices, i.e., $\mathbf{P}_i = \mathrm{diag}\left(p_{i,1}, \ldots, p_{i,L_i}\right)$. Choosing $\mathbf{P}_i$ therefore means selecting a PA policy. The power constraints on the transmitted signals $\mathbf{V}_i\mathbf{s}_i$ can be written as

$$\forall i \in \{1,2\}, \quad \mathrm{Trace}\left(\mathbf{V}_i\mathbf{P}_i\mathbf{V}_i^H\right) \leqslant M_i\, p_{i,\max}. \tag{2.3}$$

Here, we have assumed that the i.i.d. entries of matrices $\mathbf{H}_{ij}$, for all $(i,j) \in \{1,2\}^2$, are Gaussian random variables with zero mean and variance $\frac{1}{M_j}$. This assumption together with the power constraints in (2.3) is equivalent to considering a system where the entries of matrices $\mathbf{H}_{ij}$ for all $(i,j) \in \{1,2\}^2$ are Gaussian random variables with zero mean and unit variance, and the transmitted signal $\mathbf{V}_i\mathbf{s}_i$ are constrained by a finite transmit power $p_{i,\max}$. Nonetheless, the first assumption allows us to increase the dimension of the system (number of antennas) while maintaining the same average received signal to noise ratio (SNR) level $\frac{p_{i,\max}}{\sigma_i^2}$, $\forall i \in \{1,2\}$. Moreover, most of the tools from random matrix theory used in the asymptotic analysis of the achievable data rate of the opportunistic link in Sec. 2.1.3, require the variance of the entries of channel matrices to be normalized by their size. That is the reason why the normalized model, i.e., channel transfer matrices and power constraints respectively satisfying (2.1) and (2.3), was adopted.
At receiver $i \in \{1,2\}$, the signal $\mathbf{r}_i$ is processed using an $N_i \times N_i$ matrix $\mathbf{D}_i$ to form the N_i-dimensional vector $\mathbf{y}_i = \mathbf{D}_i\mathbf{r}_i$. All along this section, we refer to $\mathbf{D}_i$ as the post-processing matrix at receiver i. Regarding channel knowledge assumptions at the different nodes, we assume that the primary terminals (transmitter and receiver) have perfect knowledge of the matrix $\mathbf{H}_{11}$ while the secondary terminals have perfect knowledge of all channel transfer matrices $\mathbf{H}_{ij}$, $\forall(i,j) \in \{1,2\}^2$. One might ask whether this setup is highly demanding in terms of information assumptions. In fact, there are several technical arguments making this setup relatively realistic: (a) in some contexts channel reciprocity can be exploited to acquire CSI at the transmitters; (b) feedback channels are often available in wireless communications [110], and (c) learning mechanisms [32] can be exploited to iteratively learn the required CSI. In any case, the perfect information assumptions provide us with an upper bound on the achievable transmission rate for the secondary link.

2.1.2 Interference Alignment Strategy

In this section, we describe how both links introduced in Sec. 2.1.1 can simultaneously operate under the constraint that no additional interference is generated by the opportunistic transmitter on the primary receiver. First, we revisit the transmitting scheme implemented by the primary system [109], then we present the concept of transmit opportunity, and finally we introduce the proposed opportunistic IA technique.

Primary Link Performance

According to our initial assumptions (Sec. 2.1.1) the primary link must operate at its highest transmission rate in the absence of interference. Hence, following the results in [108, 109] and using our own notation, the optimal pre-processing and post-processing schemes for the primary link are given by the following theorem.

Theorem 2.1.1 *Let $\mathbf{H}_{11} = \mathbf{U}_{H_{11}}\mathbf{\Lambda}_{H_{11}}\mathbf{V}_{H_{11}}^H$ be a singular value decomposition (SVD) of the $N_1 \times M_1$ channel transfer matrix $\mathbf{H}_{11}$, with $\mathbf{U}_{H_{11}}$ and $\mathbf{V}_{H_{11}}$, two unitary matrices with dimension $N_1 \times N_1$ and $M_1 \times M_1$, respectively, and $\mathbf{\Lambda}_{H_{11}}$ an $N_1 \times M_1$ matrix with main diagonal $\left(\lambda_{H_{11},1}, \ldots, \lambda_{H_{11},\min(N_1,M_1)}\right)$ and zeros on its off-diagonal. The primary link achieves capacity by choosing $\mathbf{V}_1 = \mathbf{V}_{H_{11}}$, $\mathbf{D}_1 = \mathbf{U}_{H_{11}}^H$, $\mathbf{P}_1 = \mathrm{diag}(p_{1,1}, \ldots, p_{1,M_1})$, where*

$$\forall n \in \{1, \ldots, M_1\}, \quad p_{1,n} = \left(\beta - \frac{\sigma_1^2}{\lambda_{H_{11}^H H_{11},n}}\right)^+, \tag{2.4}$$

with, $\mathbf{\Lambda}_{H_{11}^H H_{11}} = \mathbf{\Lambda}_{H_{11}}^H \mathbf{\Lambda}_{H_{11}} = \mathrm{diag}\left(\lambda_{H_{11}^H H_{11},1}, \ldots, \lambda_{H_{11}^H H_{11},M_1}\right)$ and the constant β (water-level) is set to saturate the power constraint (2.3).

Let $N \triangleq \min(N_1, M_1)$. When implementing its capacity-achieving transmission scheme, the primary transmitter allocates its transmit power over an equivalent channel $\mathbf{D}_1\mathbf{H}_{11}\mathbf{V}_1 = \mathbf{\Lambda}_{H_{11}}$ which consists of at most $\mathrm{rank}(\mathbf{H}_{11}^H\mathbf{H}_{11}) \leq N$ parallel sub-channels with non-zero channel gains $\lambda_{H_{11}^H H_{11},n}$, respectively. These non-zero channel gains to which we refer as transmit dimensions, correspond to the non-zero eigenvalues of matrix $\mathbf{H}_{11}^H\mathbf{H}_{11}$. The transmit dimension $n \in \{1, \ldots, M_1\}$ is said to be used by the primary transmitter if $p_{1,n} > 0$. Interestingly, (2.4) shows that some of the transmit dimensions can be left unused. Let $m_1 \in \{1, \ldots, M_1\}$ denote the number of transmit dimensions used by the primary user:

$$\begin{aligned} m_1 &\triangleq \sum_{n=1}^{M_1} \mathbb{1}_{\{]0, M_1 p_{1,\max}]\}}(p_{1,n}) & (2.5)\\ &= \sum_{n=1}^{M_1} \mathbb{1}_{\left]\frac{\sigma_1^2}{\beta}, +\infty\right[}(\lambda_{H_{11}^H H_{11},n}). & (2.6) \end{aligned}$$

As $p_{1,\max} > 0$, the primary link transmits at least over dimension

$$n^* = \arg \max_{m \in \{1,\ldots,\min(N_1,M_1)\}} \left\{\lambda_{H_{11}^H H_{11},m}\right\} \tag{2.7}$$

regardless of its SNR, and moreover, there exist at most N transmit dimensions, thus

$$1 \leq m_1 \leq \mathrm{rank}(\mathbf{H}_{11}^H \mathbf{H}_{11}) \leq N. \tag{2.8}$$

In the following subsection, we show how those unused dimensions of the primary system can be seen by the secondary system as opportunities to transmit.

Transmit Opportunities

Once the PA matrix is set up following Th. 2.1.1, the primary equivalent channel $\mathbf{D}_1 \mathbf{H}_{11} \mathbf{V}_1 \mathbf{P}_1^{1/2} = \mathbf{\Lambda}_{H_{11}} \mathbf{P}_1^{1/2}$ is an $N_1 \times M_1$ diagonal matrix whose main diagonal contains m_1 non-zero entries and $N - m_1$ zero entries. This equivalent channel transforms the set of m_1 used and $M_1 - m_1$ unused transmit dimensions into a set of m_1 receive dimensions containing a noisy version of the primary signal, and a set of $N_1 - m_1$ unused receive dimensions containing no primary signal. The m_1 used dimensions are called primary reserved dimensions, while the remaining $N_1 - m_1$ dimensions are named secondary transmit opportunities (TO). The IA strategy, described in Section 2.1.2, allows the secondary user to exploit these $N_1 - m_1$ receive dimensions left unused by the primary link, while avoiding to interfere with the m_1 receive dimensions used by the primary link.

Definition 2.1.2 (Transmit Opportunities) *Let $\lambda_{H_{11}^H H_{11},1}, \ldots \lambda_{H_{11}^H H_{11},M_1}$ be the eigenvalues of matrix $\mathbf{H}_{11}^H \mathbf{H}_{11}$ and β be the water-level in (Th. 2.1.1). Let m_1, as defined in (2.6), be the number of primary reserved dimensions. Then the number of transmit opportunities S available to the opportunistic terminal is given by*

$$S \triangleq N_1 - m_1 = N_1 - \sum_{n=1}^{M_1} \mathbb{1}_{\left]\frac{\sigma_1^2}{\beta}, +\infty\right[}(\lambda_{H_{11}^H H_{11},n}). \tag{2.9}$$

Note that in this definition it is implicitly assumed that the number of TOs is constant over a duration equal to the channel coherence time.
Combining (2.8) and (2.9) yields the bounds on the number of transmit opportunities

$$N_1 - N \leq S \leq N_1 - 1. \tag{2.10}$$

A natural question arises as to whether the number of TOs is sufficiently high for the secondary link to achieve a significant transmission rate. In order to provide an element of response to this question, a method to find an approximation of the number of TOs per primary transmit antenna, S_∞, is proposed in Section 2.1.3. In any case, as we shall see in the next subsection, to take advantage of the TOs described here, a specific signal processing scheme is required in the secondary link.

Pre-processing Matrix

In this subsection, we define the interference alignment condition to be met by the secondary transmitter and determine a pre-processing matrix satisfying this condition.

Definition 2.1.3 (IA condition) *Let* $\mathbf{H}_{11} = \mathbf{U}_{H_{11}}\mathbf{\Lambda}_{H_{11}}\mathbf{V}_{H_{11}}^H$ *be an SVD of* $\mathbf{H}_{11}$ *and*

$$\mathbf{R} = \sigma_1^2\mathbf{I}_{N_1} + \mathbf{U}_{H_{11}}^H\mathbf{H}_{12}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{12}^H\mathbf{U}_{H_{11}}, \tag{2.11}$$

be the covariance matrix of the co-channel interference (CCI) plus noise signal in the primary link. The opportunistic link is said to satisfy the IA condition if its opportunistic transmission is such that the primary link achieves the transmission rate of the equivalent single-user system, which translates mathematically as

$$\begin{aligned} &\log_2\left|\mathbf{I}_{N_1} + \tfrac{1}{\sigma_1^2}\mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H\right| = \\ &\log_2\left|\mathbf{I}_{N_1} + \mathbf{R}^{-1}\mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H\right|. \end{aligned} \tag{2.12}$$

Our objective is first to find a pre-processing matrix $\mathbf{V}_2$ that satisfies the IA condition and then, to tune the PA matrix $\mathbf{P}_2$ and post-processing matrix $\mathbf{D}_2$ in order to maximize the transmission rate for the secondary link.

Lemma 2.1.4 (Pre-processing matrix $\mathbf{V}_2$) *Let* $\mathbf{H}_{11} = \mathbf{U}_{H_{11}}\mathbf{\Lambda}_{H_{11}}\mathbf{V}_{H_{11}}^H$ *be an ordered SVD of* $\mathbf{H}_{11}$*, with* $\mathbf{U}_{H_{11}}$ *and* $\mathbf{V}_{H_{11}}$*, two unitary matrices of size* $N_1 \times N_1$ *and* $M_1 \times M_1$*, respectively, and* $\mathbf{\Lambda}_{H_{11}}$ *an* $N_1 \times M_1$ *matrix with main diagonal* $\left(\lambda_{H_{11},1}, \ldots, \lambda_{H_{11},\min(N_1,M_1)}\right)$ *and zeros on its off-diagonal, such that* $\lambda_{H_{11},1}^2 \geqslant \lambda_{H_{11},2}^2 \geqslant \ldots \geqslant \lambda_{H_{11},\min(N_1,M_1)}^2$. *Let also the* $N_1 \times M_2$ *matrix* $\tilde{\mathbf{H}} \triangleq \mathbf{U}_{H_{11}}^H\mathbf{H}_{12}$ *have a block structure,*

$$\tilde{\mathbf{H}} = \begin{array}{r} \\ m_1 \updownarrow \\ N_1 - m_1 \updownarrow \end{array} \begin{array}{c} \overset{M_2}{\longleftrightarrow} \\ \begin{pmatrix} \tilde{\mathbf{H}}_1 \\ \tilde{\mathbf{H}}_2 \end{pmatrix} \end{array}. \tag{2.13}$$

The IA condition (Def. 2.1.3) is satisfied independently of the PA matrix $\mathbf{P}_2$*, when the pre-processing matrix* $\mathbf{V}_2$ *satisfies the condition:*

$$\tilde{\mathbf{H}}_1\mathbf{V}_2 = \mathbf{0}_{m_1 \times L_2}, \tag{2.14}$$

where L_2 is the dimension of the null space of matrix $\tilde{\mathbf{H}}_1$.

Proof: See Appendix B. □

Another solution to the IA condition was given in [76], namely $\mathbf{V}_2 = \mathbf{H}_{12}^{-1}\mathbf{U}_{H_{11}}\bar{\mathbf{P}}_1$ for a given diagonal matrix $\bar{\mathbf{P}}_1 = \mathrm{diag}\left(\bar{p}_{1,1}, \ldots, \bar{p}_{1,M_1}\right)$, with $\bar{p}_{1,n} = \left(\frac{\sigma_2^2}{\lambda_{H_{11}^H H_{11},n}} - \beta\right)^+$, where β is the water-level of the primary system (*Th. 2.1.1*) and $n \in \{1, \ldots, M_1\}$. However, such a solution is more restrictive than (2.14) as it requires $\mathbf{H}_{12}$ to be invertible and does not hold for the case when $N_i \neq M_j$, $\forall (i,j) \in \{1,2\}^2$.

Plugging $\mathbf{V}_2$ from (2.14) into (2.11) shows that to guarantee the IA condition (2.1.3), the opportunistic transmitter has to avoid interfering with the m_1 dimensions used by the primary transmitter. That is the reason why we refer to our technique as OIA: interference from the secondary user is made orthogonal to the m_1 receive dimensions used by the primary link. This is achieved by aligning the interference

from the secondary user with the $N_1 - m_1$ non-used receive dimensions of the primary link.
From Lemma 2.1.4, it appears that the L_2 columns of matrix $\mathbf{V}_2$ have to belong to the null space $\mathrm{Ker}(\tilde{\mathbf{H}}_1)$ of $\tilde{\mathbf{H}}_1$ and therefore to the space spanned by the $\dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) = M_2 - \mathrm{rank}(\tilde{\mathbf{H}}_1)$ last columns of matrix $\mathbf{V}_{\tilde{H}_1}$. Here, $\tilde{\mathbf{H}}_1 = \mathbf{U}_{\tilde{H}_1} \mathbf{\Lambda}_{\tilde{H}_1} \mathbf{V}_{\tilde{H}_1}^H$ is an SVD of $\tilde{\mathbf{H}}_{11}$with $\mathbf{U}_{\tilde{H}_1}$ and $\mathbf{V}_{\tilde{H}_1}$ being two unitary matrices with dimension m_1 and M_2, respectively, and $\mathbf{\Lambda}_{\tilde{H}_1}$ being an $m_1 \times M_2$ matrix containing the vector with main diagonal $(\lambda_{\tilde{H}_{11},1}, \ldots, \lambda_{\tilde{H}_1,\min(m_1,M_2)})$. Moreover, $\lambda^2_{\tilde{H}_{11},1} \geqslant \ldots \geqslant \lambda^2_{\tilde{H}_1,\min(m_1,M_2)}$. i.e.,

$$\mathbf{V}_2 \in \mathrm{Span}\left(\mathbf{v}_{\tilde{H}_1}^{(\mathrm{rank}(\tilde{\mathbf{H}}_1)+1)}, \ldots, \mathbf{v}_{\tilde{H}_1}^{(M_2)}\right). \tag{2.15}$$

Here, for all $i \in \{1, \ldots, M_2\}$, the column vector $\mathbf{v}_{\tilde{H}_1}$ represents the i^{th} column of matrix $\mathbf{V}_{\tilde{H}_1}$ from the left to the right.
In the following, we assume that the L_2 columns of the matrix $\mathbf{V}_2$ form an orthonormal basis of the corresponding subspace (2.15), and thus, $\mathbf{V}_2^H \mathbf{V}_2 = \mathbf{I}_{L_2}$. Moreover, recalling that $\tilde{\mathbf{H}}_1$ is of size $m_1 \times M_2$, we would like to point out that:

- When $m_1 < M_2$, $\mathrm{rank}(\tilde{\mathbf{H}}_1) \leq m_1$ and $\dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) \geq M_2 - m_1$ with equality if and only if $\tilde{\mathbf{H}}_1$ is full row-rank. This means that there are always at least $M_2 - m_1 > 0$ non-null orthogonal vectors in $\mathrm{Ker}(\tilde{\mathbf{H}}_1)$, and thus, $L_2 = \dim \mathrm{Ker}(\tilde{\mathbf{H}}_1)$. Consequently, $\mathbf{V}_2$ can always be chosen to be different from the null matrix $\mathbf{0}_{M_2 \times L_2}$.

- When, $M_2 \leqslant m_1$, $\mathrm{rank}(\tilde{\mathbf{H}}_1) \leq M_2$ and $\dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) \geq 0$, with equality if and only if $\tilde{\mathbf{H}}_1$ is full column-rank. This means that there are non-zero vectors in $\mathrm{Ker}(\tilde{\mathbf{H}}_1)$ if and only if $\tilde{\mathbf{H}}_1$ is not full column-rank. Consequently, $\mathbf{V}_2$ is a non-zero matrix if and only if $\tilde{\mathbf{H}}_1$ is not full column-rank, and again $L_2 = \dim \mathrm{Ker}(\tilde{\mathbf{H}}_1)$.

Therefore, the rank of $\mathbf{V}_2$ is given by $L_2 = \dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) \leq M_2$, and it represents the number of transmit dimensions on which the secondary transmitter can allocate power without affecting the performance of the primary user. The following lower bound on L_2 holds

$$\begin{aligned} L_2 = \dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) &= M_2 - \mathrm{rank}(\tilde{\mathbf{H}}_1) \\ &\geq M_2 - \min(M_2, m_1) \\ &= \max(0, M_2 - m_1). \end{aligned} \tag{2.16}$$

Note that by processing $\mathbf{s}_2$ with $\mathbf{V}_2$ the resulting signal $\mathbf{V}_2\mathbf{s_2}$ becomes orthogonal to the space spanned by a *subset* of m_1 rows of the cross-interference channel matrix $\tilde{\mathbf{H}} = \mathbf{U}_{H_{11}}^H \mathbf{H}_{12}$. This is the main difference between the proposed OIA technique and the classical zero-forcing beamforming (ZFBF) [72], for which the transmit signal must be orthogonal to the whole row space of matrix $\tilde{\mathbf{H}}$. In the ZFBF case, the number of transmit dimensions, on which the secondary transmitter can allocate power without affecting the performance of the primary user, is given by $L_{2,BF} = \dim \mathrm{Ker}(\tilde{\mathbf{H}}) = M_2 - \mathrm{rank}(\tilde{\mathbf{H}})$. Since $\mathrm{rank}(\tilde{\mathbf{H}}_1) \leq \mathrm{rank}(\tilde{\mathbf{H}})$, we have $L_{2,BF} \leq L_2$.

This inequality, along with the observation that $\mathrm{Ker}(\tilde{\mathbf{H}}) \subseteq \mathrm{Ker}(\tilde{\mathbf{H}}_1)$, shows that any opportunity to use a secondary transmit dimension provided by ZFBF is also provided by OIA, thus OIA outperforms ZFBF. In the next subsection we tackle the problem of optimizing the post-processing matrix $\mathbf{D}_2$ to maximize the achievable transmission rate for the opportunistic transmitter.

Post-processing Matrix

Once the pre-processing matrix $\mathbf{V}_2$ has been adapted to perform IA according to (2.15), no harmful interference impairs the primary link. However, the secondary receiver undergoes the co-channel interference (CCI) from the primary transmitter. Then, the joint effect of the CCI and noise signals can be seen as a colored Gaussian noise with covariance matrix

$$\mathbf{Q} = \mathbf{H}_{21}\mathbf{V}_{H_{11}}\mathbf{P}_1\mathbf{V}_{H_{11}}^H\mathbf{H}_{21}^H + \sigma_2^2\mathbf{I}_{N_2}. \tag{2.17}$$

We recall that the opportunistic receiver has full CSI of all channel matrices, i.e., $\mathbf{H}_{i,j}$, $\forall(i,j) \in \{1,2\}^2$. Given an input covariance matrix $\mathbf{P}_2$, the mutual information between the input $\mathbf{s}_2$ and the output $\mathbf{y}_2 = \mathbf{D}_2\mathbf{r}_2$ is

$$\begin{aligned} R_2(\mathbf{P}_2,\sigma_2^2) &= \log_2\left|\mathbf{I}_{N_2}+\mathbf{D}_2\mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H\mathbf{D}_2^H\left(\mathbf{D}_2\mathbf{Q}\mathbf{D}_2^H\right)^{-1}\right| \\ &\leqslant \log_2\left|\mathbf{I}_{N_2}+\mathbf{Q}^{-\frac{1}{2}}\mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H\mathbf{Q}^{-\frac{1}{2}}\right|, \end{aligned} \tag{2.18}$$

where equality is achieved by a whitening post-processing filter $\mathbf{D}_2 = \mathbf{Q}^{-\frac{1}{2}}$ [67]. i.e., the mutual information between the transmitted signal $\mathbf{s}_2$ and $\mathbf{r}_2$, is the same as that between $\mathbf{s}_2$ and $\mathbf{y}_2 = \mathbf{D}_2\mathbf{r}_2$. Note also that expression (2.18) is maximized by a zero-mean circularly-symmetric complex Gaussian input $\mathbf{s}_2$ [109].

Power Allocation Matrix Optimization

In this subsection, we are interested in finding the input covariance matrix $\mathbf{P}_2$ which maximizes the achievable transmission rate for the opportunistic link, $R_2(\mathbf{P}_2, \sigma_2^2)$ assuming that both matrices $\mathbf{V}_2$ and $\mathbf{D}_2$ have been set up as discussed in Sec. 2.1.2 and 2.1.2, respectively. More specifically, the problem of interest in this subsection is:

$$\begin{aligned} \max_{\mathbf{P}_2} \quad & \log_2\left|\mathbf{I}_{N_2}+\mathbf{Q}^{-\frac{1}{2}}\mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H\mathbf{Q}^{-\frac{1}{2}}\right| \\ \text{s.t.} \quad & \mathrm{Trace}\left(\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\right)\leqslant M_2 p_{2,\max}. \end{aligned} \tag{2.19}$$

Before solving the optimization problem (OP) in (2.19), we briefly describe the uniform PA scheme (UPA). The UPA policy can be very useful not only to relax some information assumptions and decrease computational complexity at the transmitter but also because it corresponds to the limit of the optimal PA policy in the high SNR regime.

Uniform Power Allocation In this case, the opportunistic transmitter does not perform any optimization on its own transmit power. It rather uniformly spreads its total power among the previously identified TOs. Thus, the PA matrix $\mathbf{P}_2$ is assumed to be of the form

$$\mathbf{P}_{2,UPA} = \gamma \mathbf{I}_{L_2}, \tag{2.20}$$

where the constant γ is chosen to saturate the transmit power constraint (2.3),

$$\gamma = \frac{M_2\, p_{2,\max}}{\mathrm{Trace}\left(\mathbf{V}_2 \mathbf{V}_2^H\right)} = \frac{M_2 p_{2,\max}}{L_2}. \tag{2.21}$$

Optimal Power Allocation Here, we tackle the OP formulated in (2.19). For this purpose, we assume that the columns of matrix $\mathbf{V}_2$ are unitary and mutually orthogonal. We define the matrix $\mathbf{K} \triangleq \mathbf{Q}^{-\frac{1}{2}} \mathbf{H}_{22} \mathbf{V}_2$, where $\mathbf{K}$ is an $N_2 \times L_2$ matrix. Let $\mathbf{K} = \mathbf{U}_K \mathbf{\Lambda}_K \mathbf{V}_K^H$ be an SVD of matrix $\mathbf{K}$, where the matrices $\mathbf{U}_K$ and $\mathbf{V}_K$ are unitary matrices with dimensions $N_2 \times N_2$ and $L_2 \times L_2$ respectively. The matrix $\mathbf{\Lambda}_K$ is an $N_2 \times L_2$ matrix with at most $\min(N_2, L_2)$ non-zero singular values on its main diagonal and zeros in its off-diagonal entries. The entries in the diagonal of the matrix $\mathbf{\Lambda}_K$ are denoted by $\lambda_{K,1}, \ldots, \lambda_{K,\min(N_2,L_2)}$. Finally, the original OP (2.19) can be rewritten as

$$\begin{array}{lll} \arg\max\limits_{\mathbf{P}_2} & \log_2 \left| \mathbf{I}_{N_2} + \mathbf{\Lambda}_K \mathbf{V}_K^H \mathbf{P}_2 \mathbf{V}_K \mathbf{\Lambda}_K^H \right| & \\ \text{s.t.} & \quad \mathrm{Trace}(\mathbf{P}_2) & = \mathrm{Trace}(\mathbf{V}_K^H \mathbf{P}_2 \mathbf{V}_K) \\ & & \leqslant M_2\, p_{2,\max}. \end{array} \tag{2.22}$$

Here, we define the square matrices of dimension L_2,

$$\tilde{\mathbf{P}}_2 \triangleq \mathbf{V}_K^H \mathbf{P}_2 \mathbf{V}_K, \tag{2.23}$$

and $\mathbf{\Lambda}_{K^H K} \triangleq \mathbf{\Lambda}_K^H \mathbf{\Lambda}_K = \mathrm{diag}\left(\lambda_{K^H K,1}, \ldots, \lambda_{K^H K,L_2}\right)$. Using the new variables $\tilde{\mathbf{P}}_2$ and $\mathbf{\Lambda}_{K^H K}$, we can write that

$$\begin{array}{rcl} \left| \mathbf{I}_{N_2} + \mathbf{\Lambda}_K \mathbf{V}_K^H \mathbf{P}_2 \mathbf{V}_K \mathbf{\Lambda}_K^H \right| & = & \left| \mathbf{I}_{L_2} + \mathbf{\Lambda}_{K^H K} \tilde{\mathbf{P}}_2 \right| \\ & \leqslant & \prod\limits_{n=1}^{L_2} \left(1 + \lambda_{K^H K,n}\, \tilde{p}_{2,n}\right) \end{array} \tag{2.24}$$

where $\tilde{p}_{2,n}$, with $n \in \{1, \ldots, L_2\}$ are the entries of the main diagonal of matrix $\tilde{\mathbf{P}}_2$. Note that in (2.24) equality holds if $\tilde{\mathbf{P}}_2$ is a diagonal matrix [54]. Thus, choosing $\tilde{\mathbf{P}}_2$ to be diagonal maximizes the transmission rate. Hence, the OP simplifies to

$$\begin{array}{ll} \max\limits_{\tilde{p}_{2,1} \cdots \tilde{p}_{2,L_2}} & \sum\limits_{n=1}^{L_2} \log_2 \left(1 + \lambda_{K^H K,n}\, \tilde{p}_{2,n}\right) \\ \text{s.t.} & \quad \sum\limits_{n=1}^{L_2} \tilde{p}_{2,n} \leqslant M_2 p_{2,\max}. \end{array} \tag{2.25}$$

The simplified optimization problem (2.25) has a water-filling solution of the form

$$\forall n \in \{1, \ldots, L_2\}, \quad \tilde{p}_{2,n} = \left(\beta_2 - \frac{1}{\lambda_{K^H K,n}}\right)^+, \tag{2.26}$$

where, the water-level β_2 is determined to saturate the power constraints in the optimization problem (2.25). Once the matrix $\tilde{\mathbf{P}}_2$ (2.23) has been obtained using water-filling (2.26), we define the optimal PA matrix $\mathbf{P}_{2,OPA}$ by

$$\mathbf{P}_{2,OPA} = \mathrm{diag}\left(\tilde{p}_{2,i}, \ldots, \tilde{p}_{2,L_2}\right), \tag{2.27}$$

while the left and right hand factors, $\mathbf{V}_K$ and $\mathbf{V}_K^H$, of matrix $\tilde{\mathbf{P}}_2$ in (2.23) are included in the pre-processing matrix:

$$\mathbf{V}_{2,OPA} = \mathbf{V}_2 \mathbf{V}_K. \tag{2.28}$$

In the next section, we study the achievable transmission rates of the opportunistic link.

2.1.3 Asymptotic Performance of the Secondary link

In this section, the performance of the secondary link is analyzed in the regime of large number of antennas, which is defined as follows:

Definition 2.1.5 (Regime of Large Numbers of Antennas) *The regime of large numbers of antennas (RLNA) is defined as follows:*

- $\forall i \in \{1, 2\}$, $N_i \to +\infty$;
- $\forall j \in \{1, 2\}$, $M_j \to +\infty$;
- $\forall (i, j) \in \{1, 2\}^2$, $\lim\limits_{\substack{M_j \to +\infty \\ N_i \to +\infty}} \frac{M_j}{N_i} = \alpha_{ij} < +\infty$, and $\alpha_{ij} > 0$ is constant.

Asymptotic Number of Transmit Opportunities

In Sec. 2.1.2, two relevant parameters regarding the performance of the opportunistic system can be identified: the number of TOs (S) and the number of transmit dimensions to which the secondary user can allocate power without affecting the performance of the primary user (L_2). Indeed, L_2 is equivalent to the number of independent symbols the opportunistic system is able to simultaneously transmit. In the following, we analyze both parameters S and L_2 in the RLNA by studying the fractions

$$S_\infty \triangleq \lim_{\substack{N_1 \to +\infty \\ M_1 \to +\infty}} \frac{S}{M_1} \text{ and,} \tag{2.29}$$

$$L_{2,\infty} \triangleq \lim_{\substack{N_1 \to +\infty \\ M_2 \to +\infty}} \frac{L_2}{M_2}. \tag{2.30}$$

Using (2.9), the fraction S_∞ can be re-written as follows

$$\begin{aligned} S_\infty &= \lim_{\substack{N_1 \to +\infty \\ M_2 \to +\infty}} \frac{1}{M_1} \left(N_1 - m_1\right) \\ &= \left(\frac{1}{\alpha_{11}} - m_{1,\infty}\right), \end{aligned} \tag{2.31}$$

where,

$$m_{1,\infty} \triangleq \lim_{\substack{N_1 \to +\infty \\ M_1 \to +\infty}} \frac{m_1}{M_1}. \tag{2.32}$$

As a preliminary step towards determining the expressions of S_∞ and $L_{2,\infty}$, we first show how to find the asymptotic water-level β_∞ in the RLNA, and the expression of $m_{1,\infty}$. First, recall from the water-filling solution (2.4) and the power constraint (2.3) that

$$\frac{1}{M_1} \sum_{n=1}^{M_1} p_{1,n} = \frac{1}{M_1} \sum_{n=1}^{M_1} \left(\beta - \frac{\sigma_1^2}{\lambda_{H_{11}^H H_{11},n}} \right)^+. \tag{2.33}$$

Define the real function q by

$$q(\lambda) = \left| \begin{array}{ll} 0, & \text{if } \lambda = 0, \\ \left(\beta - \frac{\sigma_1^2}{\lambda}\right)^+, & \text{if } \lambda > 0, \end{array} \right. \tag{2.34}$$

which is continuous and bounded on $\mathbb{R}^+$. Expression (2.33) can be rewritten as

$$\frac{1}{M_1} \sum_{n=1}^{M_1} q(\lambda_{H_{11}^H H_{11},n}) = \int_{-\infty}^{\infty} q(\lambda)\ f^{(M_1)}_{H_{11}^H H_{11}}(\lambda)\ \mathrm{d}\lambda, \tag{2.35}$$

where $f^{(M_1)}_{H_{11}^H H_{11}}$ is the probability density function associated with the empirical eigenvalue distribution $F^{(M_1)}_{H_{11}^H H_{11}}$ of matrix $\mathbf{H}_{11}^H \mathbf{H}_{11}$. In the RLNA, the empirical eigenvalue distribution $F^{(M_1)}_{H_{11}^H H_{11}}$ converges almost surely to the deterministic limiting eigenvalue distribution $F_{H_{11}^H H_{11}}$. This distribution is known as the Marčenko-Pastur law [55] whose associated density is

$$f_{H_{11}^H H_{11}}(\lambda) = \left(1 - \frac{1}{\alpha_{11}}\right)^+ \delta(\lambda) + \frac{\sqrt{(\lambda-a)^+(b-\lambda)^+}}{2\pi\lambda}, \tag{2.36}$$

where, $a = \left(1 - \frac{1}{\sqrt{\alpha_{11}}}\right)^2$ and $b = \left(1 + \frac{1}{\sqrt{\alpha_{11}}}\right)^2$. Note that the Marčenko-Pastur law has a bounded real positive support $\{\{0\} \cup [a,b]\}$ and q is continuous and bounded on $\mathbb{R}^+$. Consequently, in the RLNA, we have the almost sure convergence of (2.35), i.e.,

$$\int_{-\infty}^{\infty} q(\lambda)\ f^{(M_1)}_{H_{11}^H H_{11}}(\lambda)\ \mathrm{d}\lambda \xrightarrow{a.s.} \int_{-\infty}^{\infty} q(\lambda) f_{H_{11}^H H_{11}}(\lambda) \mathrm{d}\lambda.$$

Thus, in the RLNA (Def. 2.1.5), the water-level β_∞ is the unique solution [22] to the equation

$$\int_{\max\left(\frac{\sigma_1^2}{\beta}, a\right)}^{b} \left(\beta - \frac{\sigma_1^2}{\lambda}\right) \frac{\sqrt{(\lambda-a)(b-\lambda)}}{2\pi\lambda} \mathrm{d}\lambda - p_{1,\max} = 0, \tag{2.37}$$

and it does not depend on any specific realization of the channel transfer matrix $\mathbf{H}_{11}$, but only on the maximum power $p_{1,\max}$ and the receiver noise power σ_1^2.

We can now derive $m_{1,\infty}$. From (2.6), we have

$$\begin{aligned} m_{1,\infty} &= \lim_{\substack{N_1\to+\infty\\ M_1\to+\infty}} \frac{1}{M_1}\sum_{n=1}^{M_1} \mathbb{1}_{\left]\frac{\sigma_1^2}{\beta},+\infty\right[}(\lambda_{H_{11}^H H_{11},n}) \\ &= \lim_{\substack{N_1\to+\infty\\ M_1\to+\infty}} \int_{-\infty}^{\infty} \mathbb{1}_{\left]\frac{\sigma_1^2}{\beta},+\infty\right[}(\lambda)\ f^{(M_1)}_{H_{11}^H H_{11}}(\lambda)\ \mathrm{d}\lambda \\ &\xrightarrow{a.s.} \int_{\max(a,\frac{\sigma_1^2}{\beta_\infty})}^{b} \frac{\sqrt{(\lambda-a)(b-\lambda)}}{2\pi\lambda}\ \mathrm{d}\lambda. \end{aligned} \tag{2.38}$$

Thus, given the asymptotic number of transmist dimensions used by the primary link per primary transmit antenna $m_{1,\infty}$, we obtain the asymptotic number of transmit opportunities per primary transmit antenna S_∞ by following (2.29), i.e.,

$$S_\infty = \frac{1}{\alpha_{11}} - \int_{\max(a,\frac{\sigma_1^2}{\beta_\infty})}^{b} \frac{\sqrt{(\lambda-a)(b-\lambda)}}{2\pi\lambda}\ \mathrm{d}\lambda. \tag{2.39}$$

From (2.10), the following bounds on S_∞ hold in the RLNA:

$$\left(\frac{1}{\alpha_{11}} - 1\right)^+ \le S_\infty \le \frac{1}{\alpha_{12}}. \tag{2.40}$$

Finally, we give the expression of $L_{2,\infty}$. Recall that $L_2 = \dim \mathrm{Ker}(\tilde{\mathbf{H}}_1) = M_2 - \mathrm{rank}(\tilde{\mathbf{H}}_1)$. The rank of $\tilde{\mathbf{H}}_1$ is given by its number of non-zero singular values, or equivalently by the number of non-zero eigenvalues of matrix $\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$. Let

$$\lambda_{\tilde{H}_1^H\tilde{H}_1,1}, \ldots, \lambda_{\tilde{H}_1^H\tilde{H}_1,M_2}$$

denote the eigenvalues of matrix $\tilde{\mathbf{H}}_1^H\tilde{\mathbf{H}}_1$. We have

$$\begin{aligned} L_{2,\infty} &= 1 - \lim_{N_1,M_2\to+\infty} \frac{\mathrm{rank}(\tilde{\mathbf{H}}_1)}{M_2} \\ &= 1 - \lim_{N_1,M_2\to+\infty} \frac{1}{M_2}\sum_{n=1}^{M_2} \mathbb{1}_{]0,+\infty[}(\lambda_{\tilde{H}_1^H\tilde{H}_1,n}) \\ &= 1 - \lim_{N_1,M_2\to+\infty} \int_{-\infty}^{+\infty} \mathbb{1}_{]0,+\infty[}(\lambda) f^{(M_2)}_{\tilde{H}_1^H\tilde{H}_1}(\lambda)\mathrm{d}\lambda, \end{aligned} \tag{2.41}$$

where $f^{(M_2)}_{\tilde{H}_1^H\tilde{H}_1}(\lambda)$ is the probability density function associated with the empirical eigenvalue distribution $F^{(M_2)}_{\tilde{H}_1^H\tilde{H}_1}$. $\tilde{H}_1$ is of size $m_1 \times M_2$, and the ratio $\frac{M_2}{m_1}$ converges in the RLNA to

$$\tilde{\alpha}_1 \triangleq \lim_{N_1,M_1,M_2\to\infty} \frac{M_2}{m_1} = \frac{\alpha_{12}}{\alpha_{11} m_{1,\infty}} < \infty. \tag{2.42}$$

Thus, in the RLNA, the empirical eigenvalue distribution $F^{(M_2)}_{\tilde{H}_1^H\tilde{H}_1}$ converges almost surely to the Marčenko-Pastur law [55] $F_{\tilde{H}_1^H\tilde{H}_1}$ with associated density

$$\begin{gathered} f_{\tilde{H}_1^H\tilde{H}_1}(\lambda) = \left(1 - \frac{1}{\tilde{\alpha}_1}\right)^+ \delta(\lambda) + \frac{\sqrt{(\lambda-c)^+(d-\lambda)^+}}{2\pi\lambda}, \\ \text{where } c = \left(1 - \frac{1}{\sqrt{\tilde{\alpha}_1}}\right)^2 \text{ and } d = \left(1 + \frac{1}{\sqrt{\tilde{\alpha}_1}}\right)^2. \end{gathered} \tag{2.43}$$

Using (2.43) in (2.41) yields

$$\begin{aligned} L_{2,\infty} &\xrightarrow{a.s.} 1 - \int_{-\infty}^{+\infty} \mathbb{1}_{]0,+\infty[}(\lambda) f_{\tilde{H}_1^H \tilde{H}_1}(\lambda) \mathrm{d}\lambda \\ &= \int_{-\infty}^{+\infty} \mathbb{1}_{\{]-\infty,0]\}}(\lambda) f_{\tilde{H}_1^H \tilde{H}_1}(\lambda) \mathrm{d}\lambda \\ &= \left(1 - \frac{1}{\tilde{\alpha}_1}\right)^+ . \end{aligned} \tag{2.44}$$

Thus, given the asymptotic water-level β_∞ for the primary link, the asymptotic number of TOs per transmit antenna is given by the following expression

$$\begin{aligned} L_{2,\infty} &= \left(1 - \frac{\alpha_{11}}{\alpha_{12}} m_{1,\infty}\right)^+ \qquad (2.45) \\ &= \left(1 - \tfrac{\alpha_{11}}{\alpha_{12}} \int_{\max(a,\frac{\sigma_1^2}{\beta_\infty})}^{b} \tfrac{\sqrt{(\lambda-a)(b-\lambda)}}{2\pi\lambda}\, \mathrm{d}\lambda\right)^+ . \end{aligned}$$

Note that the number (S) of TOs as well as the number (L_2) of independent symbols that the secondary link can simultaneously transmit are basically determined by the number of antennas and the SNR of the primary system. From (2.29), it becomes clear that the higher the SNR of the primary link, the lower the number of TOs. Nonetheless, as we shall see in the numerical examples in Sec. 2.1.4, for practical values of SNR, there exist a non-zero number of TOs that the secondary can always exploit.

Asymptotic Transmission Rate of the Opportunistic Link

In this subsection, we analyze the behavior of the opportunistic rate per antenna

$$\bar{R}_2(\mathbf{P}_2,\sigma_2^2) \triangleq \tfrac{1}{N_2} \log_2 \left| \mathbf{I}_{N_2} + \mathbf{Q}^{-1} \mathbf{H}_{22} \mathbf{V}_2 \mathbf{P}_2 \mathbf{V}_2^H \mathbf{H}_{22}^H \right| \tag{2.46}$$

in the RLNA. Interestingly, this quantity can be shown to converge to a limit, the latter being independent of the realization of $\mathbf{H}_{22}$. In the present work, we essentially use this limit to conduct a performance analysis of the system under investigation but it is important to know that it can be further exploited, for instance, to prove some properties, or simplify optimization problems [25]. A key transform for analyzing quantities associated with large systems is the Stieltjes transform, which we define in App. A. By exploiting the Stieltjes transform and results from random matrix theory for large systems (See App. A), it is possible to find the limit of (2.46) in the RLNA. The corresponding result is as follows.

Proposition 2.1.1 (Asymptotic Transmission Rate) *Define the matrices*

$$\mathbf{M}_1 \triangleq \mathbf{H}_{21} \mathbf{V}_{H_{11}} \mathbf{P}_1 \mathbf{V}_{H_{11}}^H \mathbf{H}_{21}^H \tag{2.47}$$

$$\mathbf{M}_2 \triangleq \mathbf{H}_{22} \mathbf{V}_2 \mathbf{P}_2 \mathbf{V}_2^H \mathbf{H}_{22}^H \tag{2.48}$$

$$\mathbf{M} \triangleq \mathbf{M}_1 + \mathbf{M}_2, \tag{2.49}$$

and consider the system model described in Sec. 2.1.1 with primary and secondary links using the configuration $(\mathbf{V}_1,\ \mathbf{D}_1,\ \mathbf{P}_1)$ *and* $(\mathbf{V}_2,\ \mathbf{D}_2,\ \mathbf{P}_2)$ *described in Sec. 2.1.2, with* $\mathbf{P}_2$ *being any PA matrix independent from the noise level* σ_2^2*. Then, in the RLNA (Def. 2.1.5), under the assumption that* $\mathbf{P}_1$ *and* $\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H$ *have limiting eigenvalue distributions* F_{P_1} *and* $F_{V_2P_2V_2^H}$ *with compact support, the transmission rate per antenna of the opportunistic link* (Tx_2-Rx_2) *converges almost surely to*

$$\bar{R}_{2,\infty} = \frac{1}{\ln 2}\int_{\sigma_2^2}^{+\infty} G_{M_1}(-z) - G_M(-z)\,\mathrm{d}z, \tag{2.50}$$

where, $G_M(z)$ *and* $G_{M_1}(z)$ *are the Stieltjes transforms of the limiting eigenvalue distribution of matrices* $\mathbf{M}$ *and* $\mathbf{M}_1$*, respectively.* $G_M(z)$ *and* $G_{M_1}(z)$ *are obtained by solving the fixed point equations (with unique solution when* $z \in \mathbb{R}^-$ *[104]):*

$$G_{M_1}(z) = \frac{-1}{z - g(G_{M_1}(z))} \tag{2.51}$$

and

$$G_M(z) = \frac{-1}{z - g(G_M(z)) - h(G_M(z))}, \tag{2.52}$$

respectively, where the functions $g(u)$ *and* $h(u)$ *are defined as follows*

$$g(u) \triangleq \mathbb{E}\left[\frac{p_1}{1+\frac{1}{\alpha_{21}}p_1 u}\right], \tag{2.53}$$

$$h(u) \triangleq \mathbb{E}\left[\frac{p_2}{1+\frac{1}{\alpha_{22}}p_2 u}\right], \tag{2.54}$$

with the expectations in (2.53) and (2.54) taken on the random variables p_1 *and* p_2 *with distribution* F_{P_1} *and* $F_{V_2P_2V_2^H}$*, respectively.*

Proof: For the proof, see Appendix C. □

The (non-trivial) result in Prop. 2.1.1 holds for any power allocation matrix $\mathbf{P}_2$ independent of σ_2^2. In particular, the case of the uniform power allocation policy perfectly meets this assumption. This also means that it holds for the optimum PA policy in the high SNR regime. For low and medium SNRs, the authors have noticed that the matrix $\mathbf{P}_{2,OPA}$ is in general not independent of σ_2^2. This is because $\mathbf{P}_2$ is obtained from a water-filling procedure. The corresponding technical problem is not trivial and is therefore left as an extension of the present work.

2.1.4 Numerical Results

The Number S of Transmit Opportunities

As shown in (2.29), the number of TOs is a function of the number of antennas and the SNR of the primary link. In Fig. 2.2, we plot the number of TOs per transmit antenna S_∞ as a function of the SNR for different number of antennas in the receiver and transmitter of the primary link. Interestingly, even though the

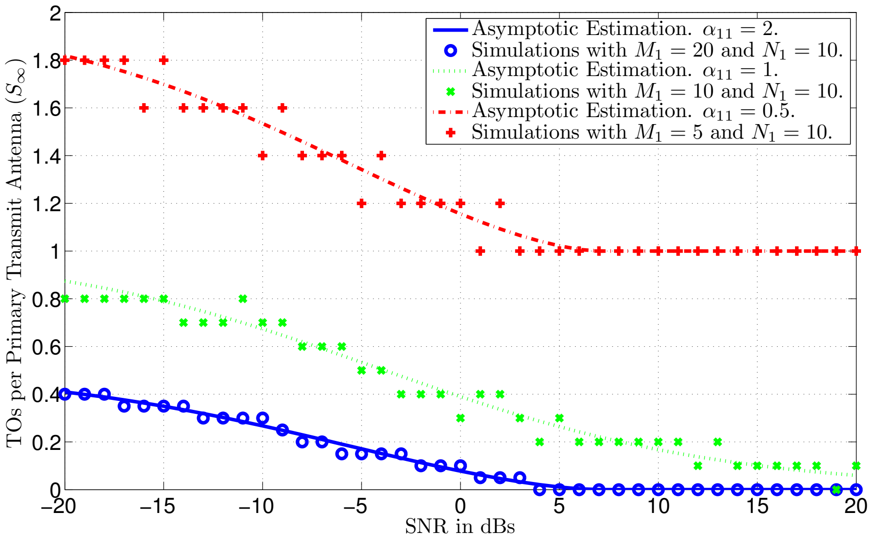

Figure 2.2: Fraction of transmit opportunities in the RLNA (Def. 2.1.5), i.e., S_∞ (2.29) as function of the SNR $= \frac{p_{1,\max}}{\sigma_1^2}$ and $\alpha_{11} = \frac{M_1}{N_1}$. Simulation results are obtained by using one realization of the matrix $\mathbf{H}_{11}$ when $N_1 = 10$.

number of TOs is a non-increasing function of the SNR, Fig. 2.2 shows that for practical values of the SNR (10 - 20 dBs.) there exists a non-zero number of TOs. Note also that the number of TOs is an increasing function of the ratio ($\alpha_{11} = \frac{M_1}{N_1}$). For instance, in the case $N_1 > M_1$, i.e., $\alpha_{11} > 1$ the secondary transmitter always sees a non-zero number of TOs independently of the SNR of the primary link, and thus, opportunistic communications are always feasible. On the contrary, when $\alpha_{11} \leqslant 1$, the feasibility of opportunistic communications depends on the SNR of the primary link.

Finally, it is important to remark that, even though the analysis of the number of TOs has been done in the RLNA (Def. 2.1.5), the model is also valid for finite number of antennas. In Fig. 2.2, we have also ploted the number of TOs observed for a given realization of the channel transfer matrix $\mathbf{H}_{11}$ when $N_1 = 10$ and $\alpha_{11} \in \{\frac{1}{2}, 1, 2\}$. Therein, it can be seen how the theoretical results from (2.29) match the simulation results.

Comparison between OIA and ZFBF

We compare our OIA scheme with the zero-forcing beamforming (ZFBF) scheme [72]. Within this scheme, the pre-processing matrix $\mathbf{V}_2$, denoted by $\mathbf{V}_{2,ZFBF}$, satisfies the condition

$$\mathbf{H}_{12}\mathbf{V}_{2,ZFBF} = \mathbf{0}_{N_r,L_2}, \tag{2.55}$$

which implies that ZFBF is feasible only in some particular cases depending on the rank of matrix $\mathbf{H}_{12}$. For instance, when $M_2 \leqslant N_1$ and $\mathbf{H}_{12}$ is full column rank, the pre-processing matrix is the null matrix, i.e., $\mathbf{V}_{2,ZFBF} = \mathbf{0}_{M_2,L_2}$ and thus, no transmission takes place. On the contrary, in the case of OIA when $M_2 \leqslant N_1$, it is still possible to opportunistically transmit with a non-null matrix $\mathbf{V}_2$ in two cases as shown in Sec. 2.1.2:

- if $m_1 < M_2$,
- or if $m_1 \geq M_2$ and $\tilde{\mathbf{H}}_1$ is not full column rank.

Another remark is that when using ZFBF, if both primary and secondary receivers come close, the opportunistic link will observe a significant power reduction since both the targeted and nulling directions become difficult to distinguish. This power reduction will be less significant in the case of OIA since it always holds that $\mathrm{rank}(\mathbf{V}_2) \geqslant \mathrm{rank}(\mathbf{V}_{2,ZFBF})$ thanks to the existence of the additional TOs. Strict equality holds only when $S = \left(\frac{1}{\alpha_{11}} - 1\right)^+$. As discussed in Sec. 2.1.2, the number of TOs (S) is independent of the position of one receiver with respect to the other. In fact, it depends on the channel realization $\mathbf{H}_{11}$ and the SNR of the primary link.
In the following, for the ease of presentation, we consider that both primary and secondary devices are equipped with the same number of antennas $N_r = N_1 = N_2$ and $N_t = M_1 = M_2$, respectively. In this scenario, we consider the cases where $N_t > N_r$ and $N_t \leqslant N_r$.

Case $N_t > N_r$ In Fig. 2.3, we consider the case where $\alpha \approx \frac{5}{4}$, with $N_r \in \{3, 9\}$. In this case, we observe that even for a small number of antennas, the OIA technique is superior to the classical ZFBF. Moreover, the higher the number of antennas, the higher the difference between the performance of both techniques. An important remark here is that, at high SNR, the performance of ZFBF and OIA is almost identical. This is basically because at high SNR, the number of TOs tends to its lower bound $N_t - N_r$ (from (2.10)), which coincides with the number of spatial directions to which ZFBF can avoid intefering. Another remark is that both UPA and OPA schemes perform identically at high SNR.

Case $N_t \leqslant N_r$ In this case, the ZFBF solution is not feasible and thus, we focus only on the OIA solution. In Fig. 2.4, we plot the transmission rate for the case where $N_r = N_t \in \{3, 6, 9\}$. We observe that at high SNR for the primary link and small number of antennas, the uniform PA performs similarly as the optimal PA. For a higher number of antennas and low SNR in the primary link, the difference between the uniform and optimal PA is significant. To show the impact of the SINR of both primary and secondary links on the opportunistic transmission rate, we present Fig.2.5. Therein, it can be seen clearly that the transmission rate in the opportunistic link is inversely proportional to the SNR level at the primary link. This is due to the lack of TOs as stated in Sec. 2.1.2. For the case when $N_r < N_t$ with strict inequality, an opportunistic transmission takes place only if $N_r - N_t \leqslant S$ and $\tilde{\mathbf{H}}_{11}$ is not full column rank. Here, the behaviour of the opportunistic transmission rate is similar to the case $N_r = N_t$ with the particularity that the opportunistic transmission rate reaches zero at a lower SNR level. As in the previous case, this is also a consequence of the number of available TOs.

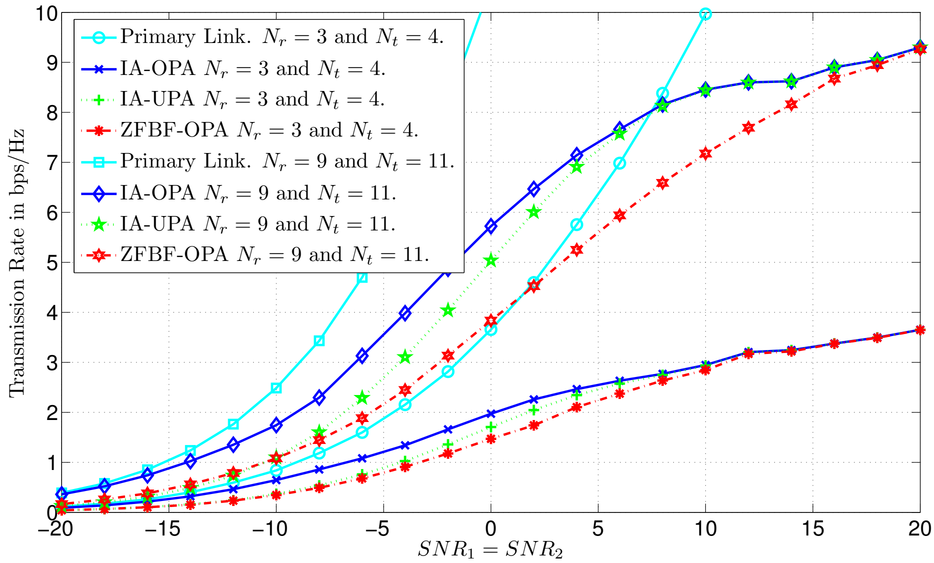

Figure 2.3: Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\mathrm{SNR}_1 = \mathrm{SNR}_2$ when IA and ZFBF are implemented. The number of antennas satisfy $\alpha = \frac{N_t}{N_r} \approx \frac{5}{4}$, with $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r \in \{3, 9\}$ and $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_1^2}$, for all $i \in \{1, 2\}$.

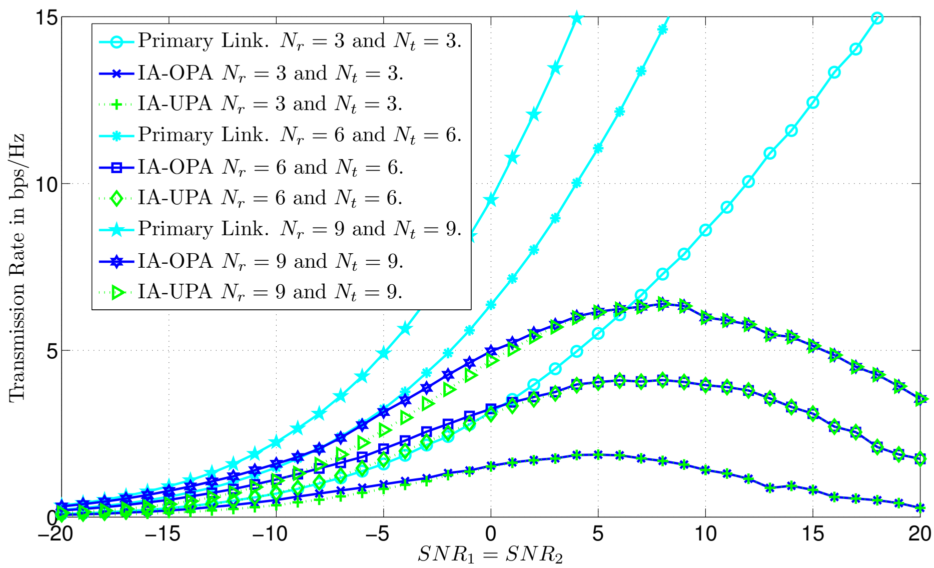

Figure 2.4: Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\mathrm{SNR}_1 = \mathrm{SNR}_2$. The number of antennas satisfy $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r$, with $N_t = N_r$, and $N_r \in \{3, 6, 9\}$ and $\mathrm{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$, for all $i \in \{1, 2\}$.

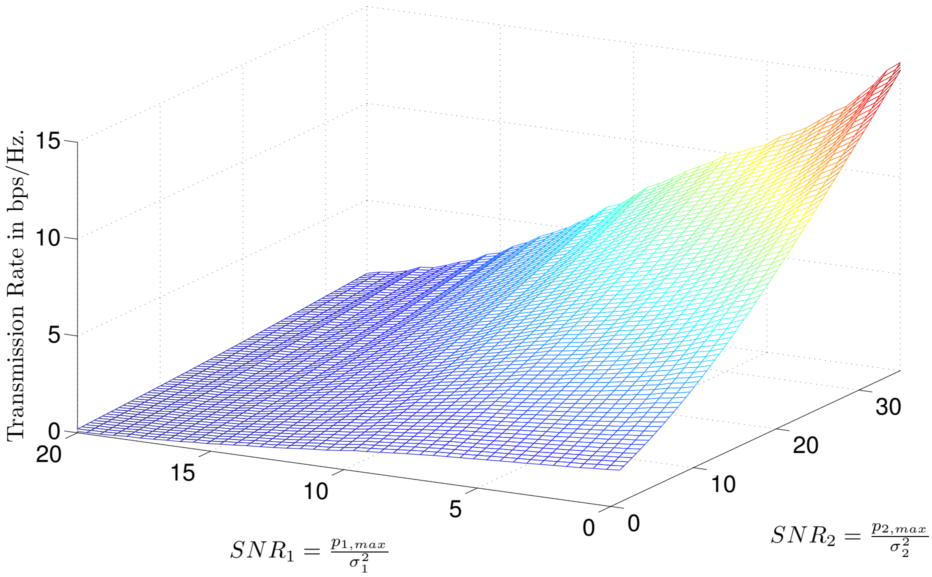

Figure 2.5: Transmission rate of the opportunistic link obtained by Monte Carlo simulations as a function of the $\text{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$, with $i \in \{1, 2\}$. The number of antennas satisfy $M_1 = M_2 = N_t$ and $N_1 = N_2 = N_r$, with $N_r = N_t = 4$.

Asymptotic Transmission Rate

In Fig. 2.6, we plot both primary and secondary transmission rates for a given realization of matrices $\mathbf{H}_{i,j}$ $\forall (i, j) \in \{1, 2\}^2$. We also plot the asymptotes obtained from Prop. 2.1.1 considering UPA in the secondary link and the optimal PA of the primary link (2.4). We observe that in both cases the transmission rate converges rapidly to the asymptotes even for a small number of antennas. This shows that Prop. 2.1.1 constitutes a good estimation of the achievable transmission rate for the secondary link even for finite number of antennas. We use Prop. 2.1.1 to compare the asymptotic transmission rates of the secondary and primary link. The asymptotic transmission rate of the primary receiver corresponds to the capacity of a single user $N_t \times N_r$ MIMO link whose asymptotes are provided in [30]. From Fig. 2.6, it becomes evident how the secondary link is able to achieve transmission rates of the same order as the primary link depending on both its own SNR and that of the primary link.

2.2 Dynamic Bandwidth Limiting Strategy

Consider now that the priority to access the spectrum is given by the order of arrival. For instance, the first to arrive can access all the available frequency bands, the second uses those left unused by the first one, and so on. Within this framework, we study the network spectral efficiency of a wireless network when the number of accessible frequency bands per transmitter is strategically limited. More specifically,

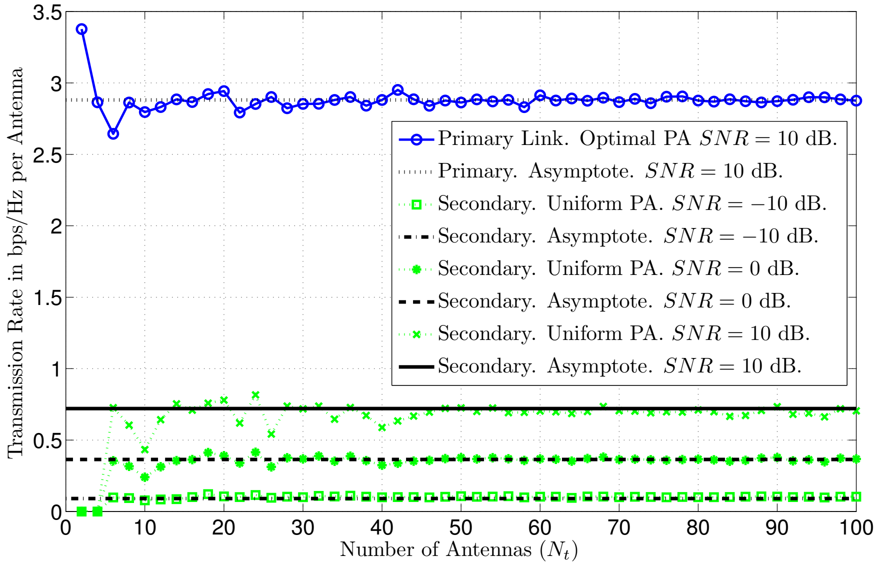

Figure 2.6: Asymptotic transmission rates per antenna of the opportunistic link as a function of the number of antennas when $N_r = N_t$ using uniform PA at different SNR levels $\text{SNR}_i = \frac{p_{i,\max}}{\sigma_i^2}$. Simulation results are obtained using one channel realization for matrices $\mathbf{H}_{ij}$ $\forall (i,j) \in \{1,2\}^2$ and theoretical results using Prop. 2.1.1,

we provide an answer to the following question: is it worth to limit the number of frequency bands each transmitter can use to improve the network spectral efficiency? And if so, how? For the ease of the presentation, we tackle these questions in the context of a parallel multiple access channel. However, the same reasoning holds for other scenarios.

In the following, particular emphasis is given to the fact that transmitters use non-intersecting sets of frequency bands (channels). That is, radio devices arriving to the network do not interfere with the existing communications. Under this condition, the optimal power allocation policy for each transmitter, regarding its individual spectral efficiency, is the water-filling power allocation over its corresponding available channels. Nonetheless, the fact that a transmitter uses several channels out of a finite set significantly reduces the total number of active transmitters. For instance, in the high signal to noise ratio (SNR) regime, few transmitters might occupy all the available channels. Following a water-filling power allocation, a given transmitter allocates the highest power levels to the channels with the highest gains. Then, the channels being used with low powers might have a negligible impact on its individual data rate. However, no other transmitter can access those channels, even though, a higher rate can be obtained by the other. This reasoning implies that there is still room for improvements in the managed spectrum access scheme, where priorities are granted by the order of arrival.

In the following, we study the benefits of limiting the number of channels each transmitter can use regarding the network spectral efficiency, i.e., the sum of all

individual spectral efficiencies. All along this section, we refer to this technique as bandwidth limiting (BL). In particular, we show the existence of an optimal BL point. Such optimal number of channels is a function of the total number of active transmitters, the total number of channels, and the different signal to noise ratios (SNR). We present simulations where we observe a significant gain in terms of spectral efficiency.
The rest of this section is structured as follows. First, the system model, which consists of a parallel multiple access channel, is introduced in Sec. 2.2.1. Then, in Sec. 2.2.2, we determine the individual and network spectral efficiency. In Sec. 2.2.3, the asymptotic network spectral efficiency is calculated and a closed form expression is obtained. In Sec. 2.2.4, we introduce the idea of bandwidth limiting. Therein, we use the asymptotic model to determine the optimal BL. Finally, numerical results are presented in Sec. 2.2.5 in order to validate our theoretical results.

2.2.1 System Model

Consider a set $\mathcal{K} = \{1, \ldots, K\}$ of transmitters communicating with a unique receiver using a set $\mathcal{N} = \{1, \ldots, N\}$ of equally spaced frequency bands (channels) as shown in Fig. 2.7. In information theory, this network topology is known as parallel MAC [23]. Transmitters arrive sequentially to the network and their index in the set $\mathcal{K}$ shows the order of arrival. All the radio devices are equipped with a unique antenna and are able to simultaneously transmit over all the channels subject to a power limitation,

$$\forall k \in \mathcal{K}, \quad \frac{1}{N}\sum_{n=1}^{N} p_{k,n} \leqslant p_{k,\max}, \tag{2.56}$$

where $p_{k,n}$ and $p_{k,\max}$ denote the transmit power over channel n and the maximum (average) transmittable power of transmitter k. In the following, we assume that all transmitters are limited by the same maximum transmittable power level, i.e., $\forall k \in \mathcal{K}$, $p_{k,\max} = p_{\max}$.

We denote the channel coefficients in the frequency domain between the receiver and transmitter k over channel n by $h_{k,n}$. We assume that for the entire transmission duration, all the channel realizations remain constant. For all $n \in \mathcal{N}$ and for all $k \in \mathcal{K}$, $h_{k,n}$ is a realization of a complex random variable h with independent and identically distributed (i.i.d) Gaussian real and imaginary parts with zero mean and variance $\frac{1}{2}$. The channel gain is denoted by $g_{k,n} = ||h_{k,n}||^2$. Then, the channel gains can be modeled by realizations of a random variable g with exponential distributions with parameter $\rho = 1$, whose cumulative distribution function (c.d.f) and probability density function (p.d.f) are denoted by $F_g(\lambda) = 1 - e^{-\lambda}$ and $f_g(\lambda) = e^{-\lambda}$, respectively. The received signals sampled at symbol rate can be written as a vector $\mathbf{y} = (y_1, \ldots, y_N)$ where the entries y_n for all $n \in \mathcal{N}$ represent the received signal over channel n. Hence,

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{s}_k + \mathbf{w}, \tag{2.57}$$

where $\mathbf{H}_k$ is an N-dimensional diagonal matrix with main diagonal $(h_{k,1}, \ldots, h_{k,N})$.

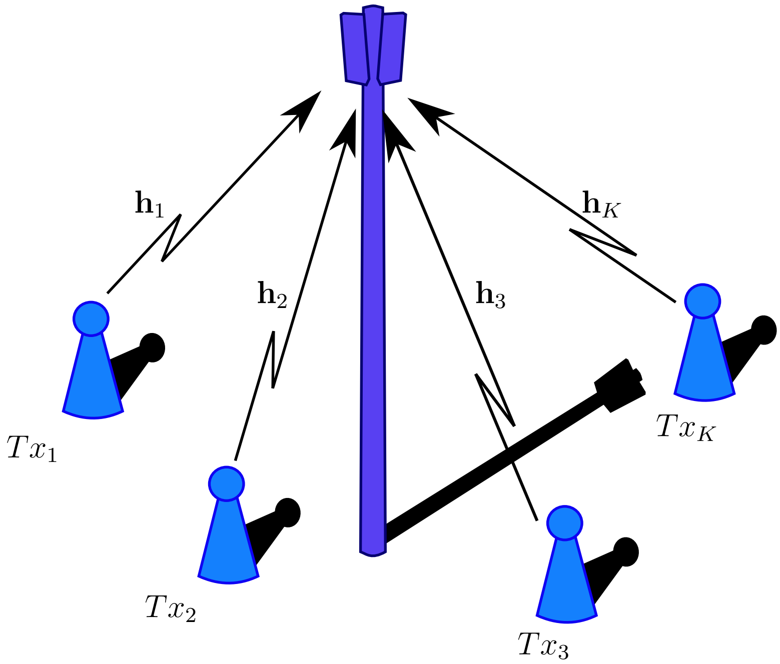

Figure 2.7: Vector multiple access channel with K transmitters and N available channels with $\mathbf{h_k} = (h_{k,1}, \ldots, h_{k,N})$ for all $k \in \mathcal{K}$.

The N-dimensional vector $\mathbf{s}_k = (s_{k,1}, \ldots, s_{k,N})$ represents the symbols transmitted by transmitter k over each channel. The power allocation profile of transmitter k, the vector $(p_{k,1}, \ldots, p_{k,N})$, is the diagonal of the diagonal matrix $\mathbf{P}_k = \mathbb{E}\left[\mathbf{s}_k \mathbf{s}_k^H\right]$. The N-dimensional vector $\mathbf{w}$ represents the noise at the receiver. Its entries, w_n for all $n \in \mathcal{N}$, are modeled by a complex circularly symmetric additive white Gaussian noise (AWGN) process with zero mean and variance σ^2.

Regarding the channel state information (CSI), we assume that each transmitter perfectly knows its own channel coefficients and the noise at each channel (no mutual interference exists). This is the case when transmitters are able to sense its environment or the receiver feeds back this parameter as a signaling message to all the transmitters.

We denote the set of channels being used by transmitter k by $\mathcal{L}_k$, i.e., $\forall k \in \mathcal{K}$ and $\forall n \in \mathcal{L}_k$, $p_{k,n} \neq 0$, and $\forall m \in \mathcal{N} \setminus \mathcal{L}_k$, $p_{k,m} = 0$. Here, a given channel cannot be used by more than one transmitter. Thus, this is equivalent to defining the sets $\mathcal{L}_k$ for all $k \in \mathcal{K}$ as a partition of the set $\mathcal{N}$, i.e.,

- $\forall (j,k) \in \mathcal{K}^2$ and $j \neq k$, $\mathcal{L}_j \cap \mathcal{L}_k = \emptyset$,
- $\forall (j,k) \in \mathcal{K}^2$ and $j \neq k$, $\mathcal{L}_j \cup \mathcal{L}_k \subseteq \mathcal{N}$.

The SINR for transmitter k over channel n is denoted by $\gamma_{k,n}$. Here, $\forall k \in \mathcal{K}$ and $\forall n \in \mathcal{N}$,

$$\gamma_{k,n} = \frac{p_{k,n} g_{k,n}}{\sigma^2}. \tag{2.58}$$

Hence, the data rate $R_k(\gamma_k)$ is,

$$R_k(\gamma_k) = \sum_{n=1}^{N} \log_2\left(1 + \gamma_{k,n}\right), \tag{2.59}$$

with $\gamma_k = (\gamma_{k,1}, \ldots, \gamma_{k,N})$.

2.2.2 Individual and Network Spectral Efficiency

Assume that each channel bandwidth is normalized to 1 Hz. A given transmitter k sets out its transmit power levels $p_{k,n}$, $\forall n \in \mathcal{N}$ by solving the optimization problem (OP)

$$\begin{array}{cl} \max\limits_{\{p_{k,n}\}_{\forall n \in \mathcal{Z}_k}} & \sum\limits_{n \in \mathcal{Z}_k} \log_2\left(1 + \gamma_{k,n}\right) \\ \text{s.t.} & \frac{1}{N} \sum\limits_{n \in \mathcal{Z}_k} p_{k,n} \leqslant p_{\max}, \end{array} \tag{2.60}$$

where, for all $k \in \mathcal{K}$, the set $\mathcal{Z}_k = \mathcal{N} \setminus \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_{k-1}$, with $\mathcal{L}_0 = \emptyset$. Thus, $\mathcal{Z}_k$ is the set of channels available for user k.

The solution to the OP in (2.60) is given in [23] and thus, we only provide the solution hereafter; $\forall k \in \mathcal{K}$ and $\forall n \in \mathcal{Z}_k$,

$$p_{k,n} = \left(\beta - \frac{\sigma^2}{g_{k,n}}\right)^+, \tag{2.61}$$

and, $\forall n \in \mathcal{N} \setminus \mathcal{Z}_k$,

$$p_{k,n} = 0. \tag{2.62}$$

The term β is a Lagrangian multiplier, known as water-level, chosen to satisfy (2.56). The transmit power levels in (2.61) can be iteratively obtained by using the water-filling algorithm described in [23]. From expression (2.61), it can be implied that $\mathcal{L}_k \subseteq \mathcal{Z}_k$.

Once the OP in (2.60) has been solved, the (average) data rate per channel of transmitter $k \in \mathcal{K}$, is

$$\bar{R}_k(\gamma_k) = \frac{1}{|\mathcal{Z}_k|} \sum_{n \in \mathcal{Z}_k} \log_2\left(1 + \gamma_{k,n}\right), \tag{2.63}$$

and then, its spectral efficiency $\Phi_k(\gamma_k)$ is

$$\Phi_k(\gamma_k) = \underbrace{\frac{|\mathcal{Z}_k|}{N}}_{\Omega_k} \bar{R}_k(\gamma_k), \tag{2.64}$$

where, Ω_k represents the fraction of spectrum accessible for transmitter k.

2.2.3 Asymptotic Network Spectral Efficiency

We define the network spectral efficiency (NSE) $\Phi(\gamma)$ as

$$\begin{aligned}\Phi(\gamma) &= \sum_{k=1}^{K} \Phi_k(\gamma_k) \\ \Phi(\gamma) &= \sum_{k=1}^{K} \Omega_k \bar{R}_k(\gamma_k), \end{aligned} \tag{2.65}$$

with $\gamma = (\gamma_1, \dots, \gamma_K)$.

In the following, we analyze the ISE in the asymptotic regime, i.e., we assume that the number of channels (N) grows to infinity. Using this result, we determine the NSE with and without bandwidth limiting. In both cases, we provide closed form expressions. Note that in this scenario, the number of transmitters K can be kept finite in the asymptotic analysis, since it does not exist any mutual interference. Thus, each transmitter can be analyzed independently. A first result on the analysis of NSE in the absence of BL for the case of spectral resource partition is presented in [33, 34, 75]. Following the same line of the analysis presented in [33], we have that in the asymptotic regime, the data rate per channel for a given transmitter k is

$$\bar{R}_k(\gamma_k) \xrightarrow{N\to\infty} \underbrace{\int_0^{\infty} \log_2\left(1 + \frac{p_k(\lambda)\lambda}{\sigma^2}\right) \mathrm{d}F_g(\lambda)}_{\bar{R}_{k,\infty}}, \tag{2.66}$$

where the functions $p_k(\lambda)$ for all $k \in \mathcal{K}$, satisfy the power constraints,

$$\int_0^{\infty} p_k(\lambda) \mathrm{d}F_g(\lambda) = p_{\max}. \tag{2.67}$$

The function $p_k(\lambda)$, $\forall k \in \mathcal{K}$, which maximizes expression (2.66) subject to expression (2.67) is also a water-filling solution, i.e.,

$$p_k(\lambda) = \left(\beta_k - \frac{\sigma^2}{\lambda}\right)^+ . \tag{2.68}$$

Note that since all the channel coefficients are drawn from the same probability distribution $f_g(\lambda)$ described in Sec. 2.2.1 and all the transmitters have the same maximum transmittable power level, we can write that $\forall k \in \mathcal{K}$, $\bar{R}_{k,\infty} = \bar{R}_{\infty}$. Hence, the water-level β_k satisfying the condition 2.68 is the same for all the transmitters. By combining expression (2.68) and (2.67), we obtain the water-level $\beta_k = \beta^*$, in the asymptotic regime by solving the equation

$$\int_{\frac{\beta^*}{\sigma^2}}^{\infty} \left(\beta^* - \frac{\sigma^2}{\lambda}\right) \mathrm{d}F_g(\lambda) - p_{\max} = 0. \tag{2.69}$$

Using the same reasoning, the fraction Ω_k, for all $k \in \mathcal{K}$, can be approximated in the asymptotic regime by

$$\Omega_{k,\infty} = \Pr\left(\beta^* < \frac{\sigma^2}{\lambda}\right) \tag{2.70}$$

$$= \int_0^{\frac{\beta^*}{\sigma^2}} \mathrm{d}F_g(\lambda) \leqslant 1, \tag{2.71}$$

which is independent of the identity of the transmitter, since all channels follow the same probability distribution. Thus, we drop the subindex k and we write,

$$\forall k \in \mathcal{K} \quad \Omega_{k,\infty} = \Omega_\infty. \tag{2.72}$$

Then, the NSE (2.65) in the asymptotic regime Φ_∞ is

$$\Phi_\infty = \sum_{i=1}^{K} (\Omega_\infty)^{i-1} \bar{R}_\infty = \frac{1-(\Omega_\infty)^K}{1-\Omega_\infty}\bar{R}_\infty. \tag{2.73}$$

2.2.4 Bandwidth Limiting Strategy

Now, we limit the number of channels each transmitter can use. When the number of accessible channels for the transmitters is limited to $L \in \mathcal{N}$ channels, the fraction of accessible spectrum $\Omega^{(\mathrm{BL})}$ for each transmitter is

$$\Omega^{(\mathrm{BL})} = \min\left\{\Pr\left(\beta^* < \frac{\sigma^2}{\lambda}\right), \frac{L}{N}\right\}. \tag{2.74}$$

Then, BL has an effect only if $\frac{L}{N} < \Pr\left(\beta^* \leqslant \frac{\sigma^2}{\lambda}\right)$. This condition is equivalent to stating that we should limit the transmitters to use a smaller number of channels than that used on the absence of BL. Hence,

$$\Omega^{(\mathrm{BL})} \leqslant \frac{L}{N}. \tag{2.75}$$

Thus, when using BL, the network spectral efficiency is,

$$\Phi^{(\mathrm{BL})}(\gamma_k) = \sum_{k=1}^{K} \Omega^{(\mathrm{BL})} \bar{R}_k(\gamma_k) \leqslant \sum_{k=1}^{K} \frac{L}{N}\bar{R}_k(\gamma_k). \tag{2.76}$$

Assume now that both K and N grow to infinity at the same rate, i.e., $N \to \infty$, and $K \to \infty$, and $\frac{N}{K} = \alpha < \infty$. Thus, the NSE using BL, denoted by $\Phi_\infty^{(\mathrm{BL})}$, in the asymptotic regime can be written as follows,

$$\begin{aligned}
\Phi_\infty^{(\mathrm{BL})} &= \lim_{N,K\to\infty} \Phi^{(\mathrm{BL})}(\gamma_k) && (2.77)\\
&= \lim_{N,K\to\infty} \sum_{k=1}^{K} \frac{L}{N}\bar{R}_k(\gamma_k) && (2.78)\\
&= \frac{1}{\alpha} L \bar{R}_\infty, && (2.79)
\end{aligned}$$

where $\bar{R}_\infty$ is given by (2.66). Now, we investigate the existence of an optimal BL point, i.e., optimal values of the fractions $\Omega^{(\mathrm{BL})}$, such that $\Phi_\infty^{(\mathrm{BL})} \geqslant \Phi_\infty$. Thus,

$$\begin{aligned}
\lim_{N,K\to\infty} \sum_{k=1}^{K} \Omega_k \bar{R}_k(\gamma_k) &\leqslant \lim_{N,K\to\infty} \sum_{k=1}^{K} \frac{L}{N}\bar{R}_k(\gamma_k)\\
\frac{1}{1-\Omega_\infty}\bar{R}_\infty &\leqslant \frac{1}{\alpha} L \bar{R}_\infty\\
L &\geqslant \frac{\alpha}{1-\Omega_\infty}. && (2.80)
\end{aligned}$$

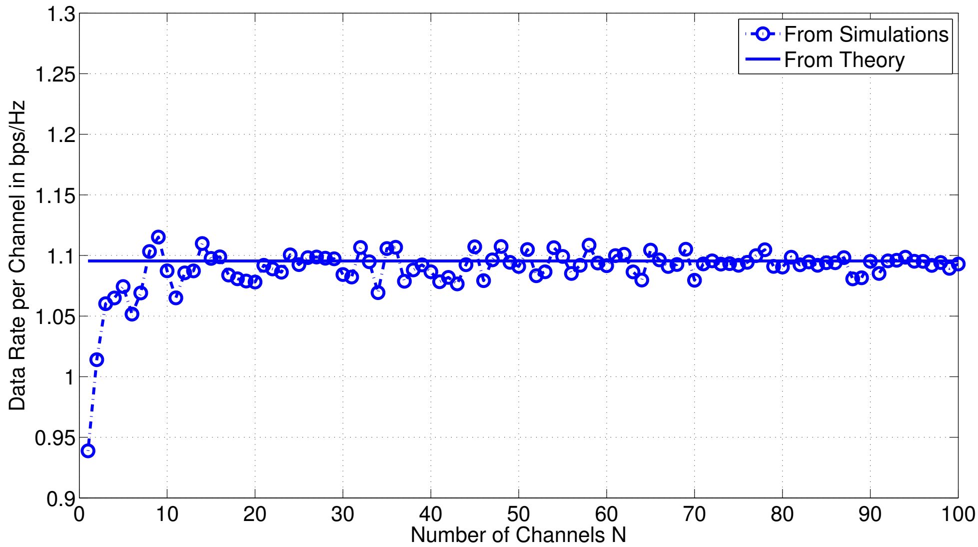

Figure 2.8: Asymptotic transmission rate per channel (2.66) in bps/Hz as a function of the number of available channels N. Dashed lines are obtained from simulations considering $\frac{p_{\max}}{\sigma^2} = 20dB$.

Thus, the optimal BL parameter, which we denote by L^* is

$$L^* = \min\left(N, \frac{\alpha}{1 - \Omega_\infty}\right). \tag{2.81}$$

In the expression above we show that the optimal BL parameter L^* depends mainly on the network load (transmitters per channel, $\alpha = \frac{K}{N}$) and the SNR of the transmitters. Note that the factor Ω_∞ is a function of $p_{\max}$, σ^2 and the probability distribution of the channels gains $f_g(\lambda)$ described in (2.70).

2.2.5 Numerical Results

In this section, we provide numerical results of our mathematical model. First, we compare the asymptotical expressions of the NSE with those obtained by simulations. In Fig. 2.8, we plot the ISE of a given transmitter. Therein, we observe that our asymptotic model (2.66) perfectly describes the system even in the finite case i.e., when K and N are small numbers. We also present simulations of the NSE obtained as a function of the BL parameter L for different network loads. In Fig. 2.9, we observe the existence of an optimum BL point. We compare the optimal BL parameter L obtained from simulations with that obtained from expression (2.80). In Fig. 2.10 we plot both results. Therein, we show that the asymptotical approximation (2.80) is a precise approximation of the optimal number of channels each transmitter must use to maximize the NSE. Finally, we show in Fig. 2.11 the NSE obtained in the absence and presence of BL. Therein, we observe a significant gain in NSE when BL is used. This gain is more important for non-overloaded networks,

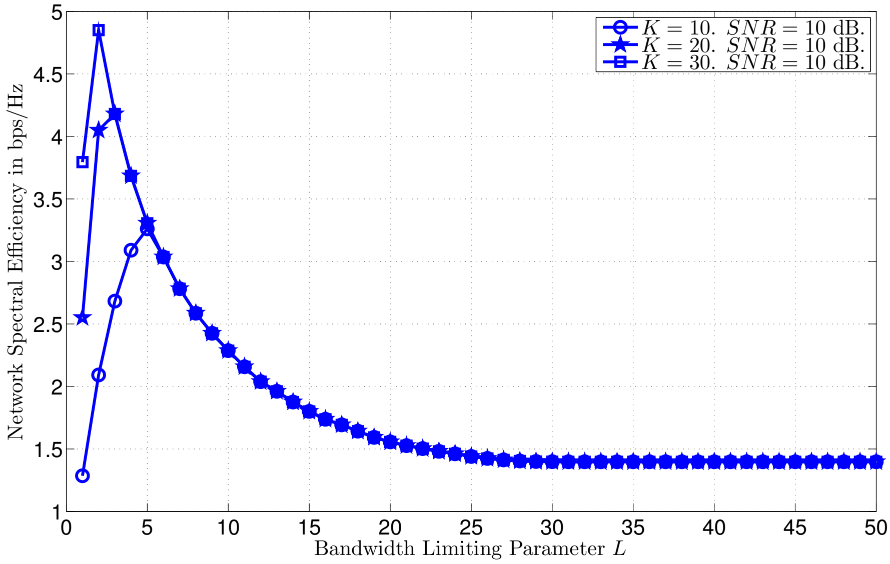

Figure 2.9: Network Spectral Efficiency (2.65) in bps/Hz as a function of the maximum number of accessible channels L. Total number of available channels $N = 50$, and $\frac{p_{\max}}{\sigma^2} = 10dB$.

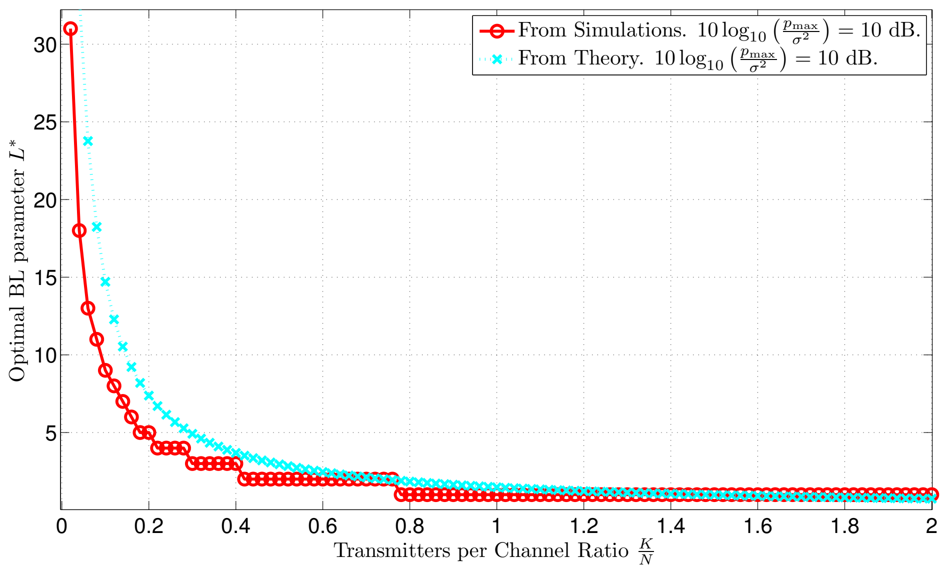

Figure 2.10: Optimal BL parameter L (2.80) as a function of the network load, $(\frac{K}{N})$. Total number of available channels $N = 50$, and $\frac{p_{\max}}{\sigma^2} = 10dB$.

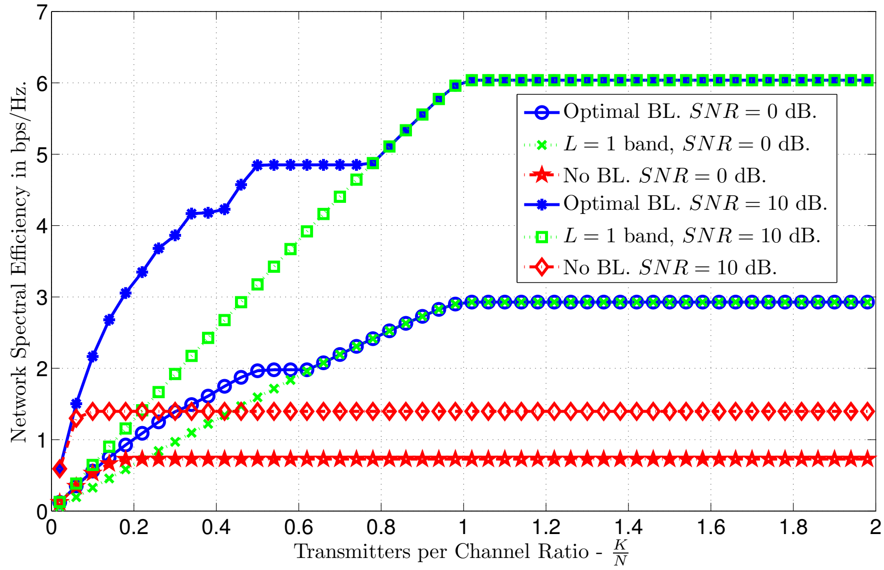

Figure 2.11: Network Spectral Efficiency (2.65) in bps/Hz as a function of the network load ($\frac{K}{N}$). Total number of available channels $N = 50$.

whereas for quasi full-loaded or overloaded networks ($K \geqslant N$), the gain obtained by BL approaches that of limiting the transmitters to use a unique channel. In the same figure, we observe that the NSE appears to be constant for certain intervals. This is due to the fact that inside those intervals the optimal BL parameter remains constant, as shown in Fig. 2.10. Moreover, the gain in NSE is quite significant at high SNR (SNR $= \frac{p_{\max}}{\sigma^2}$) levels. On the contrary, for low SNR levels, small gains in NSE are obtained when the network is low loaded.

2.3 Conclusions

In this section, two different hierarchical spectrum access scenarios were studied. In the first scenario, we consider a two-layer hierarchy. One layer for the legacy system (top priority) and another one for the opportunistic system (low priority). In the second scenario, a multi-layer hierarchy was considered. Here, the order of arrival of the radio devices determines their position in the hierarchy. For instance, the first to arrive gets the top priority and the last one the gets lowest priority for accessing the spectrum.

In the first scenario, we proposed a technique to recycle spatial directions left unused by a primary MIMO link, so that they can be re-used by secondary links. Interestingly, the number of spatial directions can be evaluated analytically, and it is shown to be sufficiently high to allow a secondary system to achieve a significant transmission rate. We provided a signal construction technique to exploit those spatial resources and a power allocation policy which maximizes the opportunistic transmission rate. Based on our asymptotical analysis, we show that this technique allows a secondary link to achieve transmission rates of the same order as those of the primary link, depending on their respective SNRs.

In the second scenario, we showed (at least in the context of a parallel multiple access channel) that the network spectral efficiency can be improved by limiting the number of available channels each transmitter can use (bandwidth limiting). We provided closed form expressions for the optimal maximum number of channels each transmitter must access in the case where transmitters use non-intersecting sets of channels. In this case, such an optimum operating point depends mainly on the network load (transmitters per channel) and the different signal to noise ratios.

Chapter 3

Equilibrium Analysis in Open Spectrum Sharing Games

In this chapter, the open spectrum access scheme described in Chapter 1 is studied using tools from game theory. In particular, this analysis is carried out considering a specific network topology, namely, the parallel multiple access channel (parallel-MAC). Multiple access channels (MAC) correspond to a scenario in which there are several transmitters and one receiver. In parallel MAC, each transmitter can exploit several orthogonal channels. This model allows the analysis of macro-diversity in the uplink of cellular networks (in this case, base stations are assumed to be connected to a common central entity), power allocation in frequency-selective multiple access channels when orthogonal frequency division multiplexing (OFDM) is used, or access point selection in wireless local area networks. In terms of multi-user channels, the channel under study corresponds to a special case of the vector Gaussian MAC [121] but here, the system is assumed to be decentralized, that is, transmitters can freely choose their spectrum access policy. This choice can be either a power allocation (PA) policy among the available channels or a channel selection (CS) policy. The performance metric for each terminal is assumed to be its individual spectral efficiency (ISE). We will refer to these problems as problem/game (a) and problem/game (b) respectively. Problems (a) and (b) can be modeled by strategic-form games where the players are the transmitters, the payoff/reward/utility function of the players is the ISE, and the set of actions consists of the possible power allocation or channel selection policies. The solution concept used in this chapter is that of Nash equilibrium (NE) [66]. One of the important reasons for making this concept relevant is that it can be the result of many evolution or dynamic processes involving reasonable information assumptions.

The most relevant existing contributions related to the problems addressed in this chapter are [6, 8, 99, 100]. The two works by Scutari et al. [99, 100] concern power allocation games in Gaussian multiple input multiple output (MIMO) interference channels while the one by Belmega et al. [6] focuses on fast fading MIMO MAC. In terms of signal model, the parallel Gaussian MAC is a special case of the Gaussian MIMO interference channels (IC) and the Gaussian MIMO MAC. However, the study of the decentralized parallel Gaussian MAC deserves some relevance due to

the following reasons.
(i) The Nash equilibrium analysis of the power allocation games in parallel MAC is not a special case of those presented in [99, 100]. In particular, the sufficient condition provided in [100] is not necessary in general [90]. The (more explicit) sufficient conditions for uniqueness given by [99] are generally not verified in parallel MAC. More precisely, from [99] it is implied that there exists a unique pure NE with high probability when for each point-to-point communication the signal dominates the interference. This condition is clearly not verified in parallel MAC. For instance, in a 2-user parallel MAC, if one user's signal is dominated by the interference, the converse holds for the other user.
(ii) The Nash equilibrium analysis of the power allocation games in fast fading MIMO MAC is not a special case of the one conducted by [6]. Indeed, in [6] the key of unconditional uniqueness is not only due to the validity of a trace inequality [5] but also to the fact that ergodic rates are considered. The latter argument does not hold in the Gaussian MIMO MAC and therefore, in parallel MAC, when channels are considered static.
(iii) The channel selection problem and corresponding (finite) game are not studied in [6, 8, 99, 100] while it is of practical interest (hard handovers in cellular networks, access point selection in wireless local area networks, etc). Although it corresponds to a special case of power allocation policy, the equilibrium analysis of this problem is different. As we shall see, there exists a Braess paradox [15]: the network sum rate or network spectral efficiency can be higher when implementing channel selection instead of a general power allocation.
(iv) In parallel MAC, when single-user decoding (SUD) is assumed at the receiver(s), both the power allocation and channel selection games have a special structure, namely, they are (best-response) potential games [115]. This attractive feature, in terms of existence of the NE and the convergence of some dynamics toward the NE [82, 86], has not been exploited in [8], where both the power allocation and channel selection games were introduced. More importantly, the games studied in [99, 100] are not potential games.
(v) A detailed study of two special cases is conducted: the $2-$transmitter $2-$channel case (small networks) and the case of large networks. The former allows one to prove the existence of a Braess paradox, while the latter allows one to predict the fraction of users using a specific channel.
The content of this chapter can be briefly summarized as follows. In Sec. 3.1, the system and game models are provided. In Sec. 3.2, it is shown that both the power allocation and channel selection games are potential games. Based on this result, the existence and uniqueness issues for Nash equilibrium in the power allocation game are studied in Sec. 3.3. The channel selection game is analyzed in Sec. 3.4. Therein, the existence of a Nash equilibrium is studied; the number of possible pure Nash equilibria is upper bounded by using an interpretation borrowed from graph theory; examples of converging dynamics are given and discussed. Sec. 3.5 provides a detailed analysis for two special cases: the $2-$transmitter$-2-$channel model and the large system model where both the numbers of transmitters and

available channels are large. The chapter is concluded by Sec. 3.7.

3.1 Models

3.1.1 Signal Model

Let us define the sets $\mathcal{K} \triangleq \{1, \ldots, K\}$ and $\mathcal{S} \triangleq \{1, \ldots, S\}$. Consider a parallel multiple access channel with K transmitters and S subchannels (namely non-overlapping bands). Denote by $\mathbf{y} = (y_1, \ldots, y_S)^T$ the S-dimensional vector representing the received signal, which can be written in the baseband at the symbol rate as follows

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{x}_k + \mathbf{w}. \tag{3.1}$$

Here, $\forall k \in \mathcal{K}$, $\mathbf{H}_k$ is the channel transfer matrix from transmitter k to the receiver, $\mathbf{x}_k$ is the vector of symbols transmitted by transmitter k, and vector $\mathbf{w}$ represents the noise observed at the receiver. We will exclusively deal with the scenario where $\forall k \in \mathcal{K}$, matrix $\mathbf{H}_k$ is an S-dimensional diagonal matrix (parallel MAC), i.e., $\mathbf{H}_k = \mathrm{diag}\,(h_{k,1}, \ldots, h_{k,S})$. In our analysis, block fading channels are assumed. Hence, for each channel use, the entries $h_{k,s}$, for all $(k,s) \in \mathcal{K} \times \mathcal{S}$, are time-invariant realizations of a complex circularly symmetric Gaussian random variable with zero mean and unit variance. Here, it is also assumed that each transmitter $k \in \mathcal{K}$ is able to perfectly estimate its own channel realizations (coherent communications), i.e., the channels $h_{k,1} \ldots h_{k,S}$. The vector of transmitted symbols $\mathbf{x}_k$, $\forall k \in \mathcal{K}$, is an S-dimensional complex circularly symmetric Gaussian random variable with zero mean and covariance matrix $\mathbf{P}_k = \mathbb{E}\left(\mathbf{x}_k \mathbf{x}_k^H\right) = \mathrm{diag}\,(p_{k,1}, \ldots, p_{k,S})$. Assuming the input data flows to be Gaussian and independent is optimal in terms of spectral efficiency, as shown in [108, 109]. For all $(k,s) \in \mathcal{K} \times \mathcal{S}$, $p_{k,s}$ represents the transmit power allocated by transmitter k over channel s. Transmitters are power-limited, that is,

$$\forall k \in \mathcal{K}, \quad \sum_{s=1}^{S} p_{k,s} \leqslant p_{k,\max}, \tag{3.2}$$

where $p_{k,\max}$ is the maximum transmit power of transmitter k. A power allocation (PA) vector for transmitter $k \in \mathcal{K}$ is any vector $\mathbf{p}_k = (p_{k,1}, \ldots, p_{k,S})$ with non-negative entries satisfying (3.2). The noise vector $\mathbf{w}$ is an S-dimensional zero mean Gaussian random variable with independent, equal variance real and imaginary parts. Here, $\mathbb{E}\left(\mathbf{w}\mathbf{w}^H\right) = \mathrm{diag}\,(\sigma_1^2, \ldots, \sigma_S^2)$, where, σ_s^2 represents the noise power over channel s. We respectively denote the noise spectral density and the bandwidth of channel $s \in \mathcal{S}$ by N_0 and B_s, and thus, $\sigma_s^2 = N_0 B_s$. The total bandwidth is denoted by $B = \sum_{s=1}^{S} B_s$.

3.1.2 Game Model

The PA and CS problems described above can be respectively modeled by the following two non-cooperative static games in strategic form (with $i \in \{a, b\}$):

$$\mathcal{G}_{(i)} = \left(\mathcal{K}, \left(\mathcal{P}_k^{(i)}\right)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right). \tag{3.3}$$

In both games, the set of transmitters $\mathcal{K}$ is the set of players. An action of a given transmitter $k \in \mathcal{K}$ is a particular PA scheme, i.e., an S-dimensional PA vector $\mathbf{p}_k = (p_{k,1}, \ldots, p_{k,S}) \in \mathcal{P}_k^{(i)}$, where $\mathcal{P}_k^{(i)}$ is the set of all possible PA vectors which transmitter k can use either in the game $\mathcal{G}_{(a)}$ $(i = a)$ or in the game $\mathcal{G}_{(b)}$ $(i = b)$. An action profile of the game $i \in \{a, b\}$ is a super vector

$$\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_K) \in \mathcal{P}^{(i)},$$

where $\mathcal{P}^{(i)}$ is a set obtained from the Cartesian product of the action sets $\mathcal{P}_k^{(i)}$, for all $k \in \mathcal{K}$, i.e., $\mathcal{P}^{(i)} = \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_K^{(i)}$, where,

$$\begin{aligned}
\mathcal{P}_k^{(a)} &= \Big\{(p_{k,1}, \ldots p_{k,S}) \in \mathbb{R}^S : \forall s \in \mathcal{S}, p_{k,s} \geqslant 0, \\
&\qquad \sum_{s\in\mathcal{S}} p_{k,s} \leqslant p_{k,\max}\Big\}, \text{ and} &(3.4)\\
\mathcal{P}_k^{(b)} &= \{p_{k,\max}\, \mathbf{e}_s : \forall s \in \mathcal{S},\ \mathbf{e}_s = (e_{s,1}, \ldots, e_{s,S}), \\
&\qquad \forall r \in \mathcal{S} \setminus s, e_{s,r} = 0, \text{ and } e_{s,s} = 1\}. &(3.5)
\end{aligned}$$

In the sequel, we respectively refer to the games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ as the PA game and CS game. Let us denote by $\mathbf{p}_{-k}$ any vector in the set

$$\mathcal{P}_{-k}^{(i)} \triangleq \mathcal{P}_1^{(i)} \times \ldots \times \mathcal{P}_{k-1}^{(i)} \times \mathcal{P}_{k+1}^{(i)} \times \ldots \times \mathcal{P}_K^{(i)} \tag{3.6}$$

with $(i, k) \in \{a, b\} \times \mathcal{K}$. For a given $k \in \mathcal{K}$, the vector denoted by $\mathbf{p}_{-k}$ represents the strategies adopted by all the players other than player k. With a slight abuse of notation, sometimes we write any vector $\mathbf{p} \in \mathcal{P}^{(i)}$ as $(\mathbf{p}_k, \mathbf{p}_{-k})$, in order to emphasize the k-th vector component of the super vector $\mathbf{p}$. The utility for player k in the game $\mathcal{G}_{(i)}$ is its spectral efficiency $u_k : \mathcal{P}^{(i)} \to \mathbb{R}_+$, and

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) = \sum_{s\in\mathcal{S}} \frac{B_s}{B} \log_2(1 + \gamma_{k,s}) \text{ [bps/Hz]} \tag{3.7}$$

where $\gamma_{k,s}$ is the signal-to-interference plus noise ratio (SINR) seen by player k over its channel s, i.e.,

$$\gamma_{k,s} = \frac{p_{k,s} g_{k,s}}{\sigma_s^2 + \sum\limits_{j\in\mathcal{K}\setminus\{k\}} p_{j,s} g_{j,s}}, \tag{3.8}$$

and $g_{k,s} \triangleq |h_{k,s}|^2$. Note that we assume that single-user decoding (SUD) is used at the receiver(s). Clearly, optimality is not sought here. Rather, a scalable (in terms of signaling cost) and fair choice for the decoding scheme is made. Additionally, this choice allows the games under investigation to be potential games [64], which is a very attractive feature to implement dynamic procedures aiming at converging to a Nash equilibrium.

3.2 A Note on the Parallel Multiple Access Channel

The games $\mathcal{G}_{(i)}$, $i \in \{a, b\}$, correspond to a conflict of interest between selfish decision-makers. Here, interaction is due to multiple access interference and decisions consist of the choice of the PA vectors. Based on certain information/behavior assumptions at/for the transmitters, a natural question is to know whether this conflict has some predictable outcomes. Thus, following this reasoning, we focus on the Nash equilibrium (NE) [66] as a solution concept of this conflict. Pure NE are defined as follows.

Definition 3.2.1 (Pure Nash Equilibrium) *In the non-cooperative games in strategic form $\mathcal{G}_{(i)}$, with $i \in \{a, b\}$, an action profile $\boldsymbol{p} \in \mathcal{P}^{(i)}$ is a pure NE if it satisfies, for all $k \in \mathcal{K}$ and for all $\mathbf{p}_k' \in \mathcal{P}_k^{(i)}$, that*

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) \geqslant u_k(\mathbf{p}_k', \mathbf{p}_{-k}). \tag{3.9}$$

When at least one NE exists in the games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$, it can be achieved as the result of long-term interaction of the players following a particular behavioral rule (see Sec. 3.3 and 3.4) with reduced feedback (receiver - transmitter).
Potential games (PG) [64] is a class of games for which existence of pure NE is guaranteed. Additionally, many known learning procedures, such as, best response dynamics, fictitious play and some reinforcement learning dynamics converge in PG. One of the purposes of this section is to show that the games $\mathcal{G}_{(i)}$, $i \in \{a, b\}$ can be checked to be potential games [64, 96] and more generally, best response PG (BRPG) [115]. First, let us define an exact potential game.

Definition 3.2.2 (Exact Potential Game) *Any game in strategic form defined by the triplet $\left(\mathcal{K}, (\mathcal{P}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right)$ is an exact potential game if there exists a function $\phi(\mathbf{p})$ for all $\mathbf{p} \in \mathcal{P} = \mathcal{P}_1 \times \ldots \times \mathcal{P}_K$ such that for all players $k \in \mathcal{K}$ and for all $\mathbf{p}_k' \in \mathcal{P}_k$, it holds that*

$$u_k(\mathbf{p}_k, \mathbf{p}_{-k}) - u_k(\mathbf{p}_k', \mathbf{p}_{-k}) = \phi(\mathbf{p}_k, \mathbf{p}_{-k}) - \phi(\mathbf{p}_k', \mathbf{p}_{-k}).$$

From the definition of the utility functions (3.7), the following proposition can be easily shown [73].

Proposition 3.2.1 *The strategic form games $\mathcal{G}_{(i)}$, with $i \in \{a, b\}$, are exact potential games with potential function*

$$\phi(\mathbf{p}) = \sum_{s\in\mathcal{S}} \frac{B_s}{B} \log_2 \left(\sigma_s^2 + \sum_{k=1}^{K} p_{k,s} g_{k,s} \right). \tag{3.10}$$

In fact, the games $\mathcal{G}_{(i)}$, $i \in \{a, b\}$, are not only potential games but also best-response potential games [115]. A game in normal form $\left(\mathcal{K}, (\mathcal{P}_k)_{k\in\mathcal{K}}, (u_k)_{k\in\mathcal{K}}\right)$ is a BRPG if there exists a function $\theta : \mathcal{P}_1 \times \ldots \times \mathcal{P}_K \to \mathbb{R}$, which verifies, for all $k \in \mathcal{K}$,

$$\arg \max_{\mathbf{q}_k \in \mathcal{P}_k^{(i)}} u_k\left(\mathbf{q}_k, \mathbf{p}_{-k}\right) = \arg \max_{\mathbf{q}_k \in \mathcal{P}_k^{(i)}} \theta\left(\mathbf{q}_k, \mathbf{p}_{-k}\right). \tag{3.11}$$

Indeed, the utility function (3.7) can be written in terms of the potential function (3.10) as follows

$$\forall k \in \mathcal{K}, \quad u_k\left(\mathbf{p}_k, \mathbf{p}_{-k}\right) = \phi\left(\mathbf{p}_k, \mathbf{p}_{-k}\right) - \upsilon_k\left(\mathbf{p}_{-k}\right), \tag{3.12}$$

where $\upsilon_k\left(\mathbf{p}_{-k}\right) = \sum_{s \in \mathcal{S}} \frac{B_s}{B} \log_2\left(\sigma_s^2 + \sum_{j \in \mathcal{K}\setminus\{k\}} p_{j,s} g_{j,s}\right)$, and by inspection, it becomes clear that both $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ satisfy the condition (3.11). The technical point to be noticed here is that in general an NE point does not necessarily maximizes the potential function (the converse is always true, see e.g., [97]) but here, due to (3.12), the set of NE and the set of maximum of the potential coincide. This result could also be proved alternatively by using [97] since the potential game is also concave in the sense of Theorem 1 in [92]. Thus, the individual utility maximization problem is *equivalent* to maximizing a common function (independently of the transmitter index). This reasoning leads to the following proposition.

Proposition 3.2.2 (Lemma 2.2 **in [63])** *Denote by $\mathcal{N}_{(i)}$ and $\mathcal{N}'_{(i)}$ the sets of pure NE for the games $\mathcal{G}_{(i)}$ and $\mathcal{G}'_{(i)} = \left(\mathcal{K}, \left(\mathcal{P}_k\right)_{k\in\mathcal{K}}, \left(\phi\right)_{k\in\mathcal{K}}\right)$, with $i \in \{a, b\}$. Then, for all $i \in \{a, b\}$, the sets $\mathcal{N}_{(i)}$ and $\mathcal{N}'_{(i)}$ are identical.*

Prop. 3.2.1 and Prop. 3.2.2 will be exploited in Sec. 3.3 and Sec. 3.4 to solve the pure NE existence and uniqueness problems.

3.3 The Power Allocation Game $\mathcal{G}_{(a)}$

The PA game $\mathcal{G}_{(a)}$ models the scenario where transmitters can allocate any power level to any of its own channels subject to the power constraints (3.2). First, the NE existence and uniqueness problems are tackled. Second, the PA game is considered to be played several times and the convergence issue for two known dynamics (namely the sequential and simultaneous BR dynamics) under partial information on the game is discussed.

3.3.1 Existence of a Pure NE

Our main interest is to find NE in pure strategies, i.e., a PA vector which will be played once and for all by the transmitters during a time window shorter than the channel coherence time (the PA policies have to be updated for each channel realization). First of all, note that the existence of a mixed NE (i.e., a probability distribution on the possible actions which verifies Definition 3.2.1) is guaranteed. This is because the action spaces, $\mathcal{P}_k^{(a)}$ for all $k \in \mathcal{K}$ are compact spaces and the utility functions are continuous with respect to the action profile [28]. However, in compact strategy spaces, mixed strategies are generally less attractive in the wireless communications (e.g. think of how to implement such strategies). Interestingly, the PG property allows one to assert the existence of at least one pure NE. The following proposition is an immediate consequence of *Lemma* 4.3 in [64].

Proposition 3.3.1 (Existence of a pure NE) *The game $\mathcal{G}_{(a)}$ has always at least one NE in pure strategies.*

Here, it is important to remark that more general concepts of equilibrium, such as correlated equilibrium [3], can coincide with the NE in some PG. This is the case in games where the potential is a strictly concave function [69]. The potential function (3.10) is concave but not strictly concave. Strict concavity would be obtained by considering ergodic transmission rates, which is relevant in fast fading parallel MAC.

3.3.2 Uniqueness of the Pure NE

As previously mentioned, the equilibrium analysis of the game $\mathcal{G}_{(a)}$ is not a special case of the studies conducted in [99] [100] [6]. In [99, 100] sufficient conditions are assumed to ensure uniqueness in Gaussian MIMO IC whereas in [6] ergodicity is exploited in fast fading MIMO MAC to have unconditional uniqueness. In the case of the Gaussian parallel MAC, uniqueness is guaranteed almost surely and the argument used to prove uniqueness is based on the concept of degeneracy which allows one to characterize the directions along which the potential remains constant. The following result is due to a recent work from Mertikopoulos et al. [59].

Theorem 3.3.1 (NE uniqueness in parallel MAC) The game $\mathcal{G}_{(a)}$ has almost surely a unique pure NE.

A formal proof of Theorem 3.3.1 is provided in [59]. Here, we provide a corollary of Theorem 3.3.1 which can be proved without using the concept of degeneracy. This result will be used later in Sec. 3.5.1.

Corollary 3.3.2 (Uniqueness in 2×2 systems) Let $(K, S) = (2, 2)$. Then, the game $\mathcal{G}_{(a)}$ has a unique pure NE with probability 1.

The 2×2 power allocation game is analyzed later in this chapter in Sec. 3.5.1. Here, we use Prop. 3.5.1 from Sec. 3.5.1 to give a proof of Corollary 3.3.2. Based on Prop. 3.5.1, it can be implied that a necessary (but not sufficient) condition for observing multiple NE in the game $\mathcal{G}_{(a)}$ is that the vector of channel realizations $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{G}$ satisfies $g_{11}g_{22} = g_{21}g_{12}$. However, considering that the channel gains $\{g_{i,j}\}_{\forall (i,j) \in \mathcal{K}\times\mathcal{P}}$ are drawn from continuous probability distributions, observing a channel realization such that $g_{11}g_{22} = g_{12}g_{21}$ is a zero-probability event. Thus, we conclude that the NE of the game $\mathcal{G}_{(a)}$ is unique with probability 1.
In general, from Prop. 3.2.2 and Def. 3.2.1, the (unique pure) NE of the game $\mathcal{G}_{(a)}$, denoted by $\mathbf{p}^{\dagger} = \left(\mathbf{p}_1^{\dagger}, \dots, \mathbf{p}_K^{\dagger}\right)$ is the unique solution of the following optimization problem:

$$\forall k \in \mathcal{K}, \quad \mathbf{p}_k^{\dagger} \in \arg \max_{\mathbf{p}_k \in \mathcal{P}_k^{(a)}} \phi\left(\mathbf{p}_k, \mathbf{p}_{-k}^{\dagger}\right). \tag{3.13}$$

The components of the vector $\mathbf{p}^{\dagger}$ in (3.13) are for all $(k, s) \in \mathcal{K} \times \mathcal{S}$,

$$p^{\dagger}_{k,s} = \left[\frac{B_s}{B}\frac{1}{\beta_k} - \frac{\sigma_s^2 + \sum\limits_{j\in\mathcal{K}\setminus\{k\}} p^{\dagger}_{j,s} g_{j,s}}{g_{k,s}} \right]^+ , \tag{3.14}$$

where, β_k is a Lagrangian multiplier chosen to saturate the power constraints (3.2). In the wireless communications domain, this particular power allocation scheme is known as water-filling and the constant β_k is known as the water level [121].

3.3.3 Dynamics Arising from the Best Response

From the preceding section, we know that there is a unique pure NE in $\mathcal{G}_{(a)}$. This allows one to predict the unique outcome of the static game $\mathcal{G}_{(a)}$ with complete information or for some scenarios where the game is played several times but with partial information. In a cognitive radio setting, it is typically more realistic to assume that a transmitter can sense the actions played by the others, react to it, the others sense this, update their actions, and so on. This is the idea of the best response dynamics (BRD). As dynamics come into play, the question is to know whether these dynamics converge, which is the purpose of this section.
Let us define the best response correspondence and best response dynamics.

Definition 3.3.3 (Best-Response Correspondence) *In a non-cooperative game described by the* 3*-tuple* $\left(\mathcal{K}, (\mathcal{P}_k)_{\forall k\in\mathcal{K}}, (u_k)_{\forall k\in\mathcal{K}}\right)$*, the relation* $\mathrm{BR}_k : \mathcal{P}_{-k} \to \mathcal{P}_k$ *such that*

$$\mathrm{BR}_k(\mathbf{p}_{-k}) = \arg\max_{\mathbf{q}_k\in\mathcal{P}_k} u_k(\mathbf{q}_k, \mathbf{p}_{-k}), \tag{3.15}$$

is defined as the best-response correspondence of player $k \in \mathcal{K}$*, given the actions* $\mathbf{p}_{-k}$ *adopted by all the other players.*

Definition 3.3.4 (Best Response Dynamics) *Let the action profile*

$$\mathbf{p}(t) = (\mathbf{p}_1(t), \dots, \mathbf{p}_K(t))$$

be the result of a best-response dynamics at time t*. Then, for all* $k \in \mathcal{K}$*, and for all* $t \in \mathbb{N}$*, the vector* $\mathbf{p}_k(t)$ *can be obtained as follows: (1) In the sequential best-response dynamics (round-Robin order):*

$$\mathbf{p}_k(t) \in \mathrm{BR}_k(\mathbf{p}_1(t), \dots, \mathbf{p}_{k-1}(t), \mathbf{p}_{k+1}(t-1), \dots, \mathbf{p}_K(t-1)), \tag{3.16}$$

(2) in the simultaneous best-response dynamics:

$$\mathbf{p}_k(t) \in \mathrm{BR}_k(\mathbf{p}_{-k}(t-1)), \tag{3.17}$$

where $\mathbf{p}(0)$ *can be any vector* $\mathbf{p} \in \mathcal{P}$*.*

Now, based on both definitions above and Theorem 6.2 in [120], the following result can be stated.

Proposition 3.3.2 (Convergence of BRD) *The sequential best response dynamics of any potential game converges to an NE.*

Assuming that each transmitter knows its actual channel gains $\mathbf{g}_k = (g_{k,1}, \ldots, g_{k,S})$, the bandwidth of all channels $\mathbf{b} = (B_1, \ldots, B_S)$, and its own actual PA vector $\mathbf{p}_k$, which is a realistic assumption, each transmitter can determine its best response (3.17) based on a common message from the receiver (e.g. the multiple access interference value on each channel, $\kappa = (\kappa_1, \ldots, \kappa_S)$, with $\kappa_s = \sigma_s^2 + \sum_{k=1}^{K} p_{k,s} g_{k,s}$). Note that the sequential BRD in the game $\mathcal{G}_{(a)}$ leads to the same result as the iterative water-filling algorithm (IWFA) presented in [121]. However, the IWFA has been obtained in a context where the receiver performs successive interference cancellation (SIC) and the objective was to maximize the sum rate of the network. In this work, single-user decoding is considered and each transmitter aims to maximize its own data rate. Hence, even the mathematical expressions are similar the concepts do not necessarily have the same meaning and the achieved sum rates in each case are different from each other.
One of the main drawbacks of the iterative BRD (and IWFA) is its large time for convergence as well as its required signaling (message $\kappa(t)$). To overcome this problem, other algorithms such as the simultaneous water-filling algorithm (SWFA), which follows the simultaneous best-response dynamics (Def. 3.3.4) have been proposed [101].

3.4 The Channel Selection Game $\mathcal{G}_{(b)}$

In this section, a constraint is imposed to the transmitters: they can only transmit over one channel at a time. Considering this constraint, it is easy to check that every transmitter has to saturate its transmit power to maximize its utility. Indeed, $\forall k \in \mathcal{K}$ and $\forall \mathbf{p}_{-k} \in \mathcal{P}_{-k}^{(b)}$, $\phi\left(p_k\, \mathbf{e}_s, \mathbf{p}_{-k}\right) < \phi\left(p_{k,\max}\, \mathbf{e}_s, \mathbf{p}_{-k}\right)$ where $\mathbf{e}_s = (e_{s,1}, \ldots, e_{s,S}) \in \mathbb{R}^S$, $\forall r \in \mathcal{S} \setminus s$, $e_{s,r} = 0$, and $e_{s,s} = 1$. The problem under investigation is therefore a channel selection problem. Technically, the main difference between $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ is that the latter is a finite game ($|\mathcal{K} \times \mathcal{S}| < +\infty$). As a consequence, the number of pure NE is generally more than 1.

3.4.1 Existence of a Pure NE

The game $\mathcal{G}_{(b)}$ is an exact potential game (Prop. 3.2.1) and thus, following *Lemma 2.3* in [64], we introduce the following proposition.

Proposition 3.4.1 (Existence of a pure NE) *The game $\mathcal{G}_{(b)}$ has always at least one NE in pure strategies.*

Here again, even though the focus is on pure strategy NE, it is useful to remark that given that the actions sets are discrete and finite, then the existence of at least one NE in mixed strategies is ensured [66]. In the case of 2 players and 2 channels, the NE in mixed-strategies has been analyzed in [86].

3.4.2 Bounding the Number of Pure NE

To identify an upper-bound of the number of NE in $\mathcal{G}_{(b)}$, we exploit basic tools from graph theory. Let us index the elements of the action set $\mathcal{P}^{(b)}$ in any given order using the index $n \in \mathcal{I} = \left\{1, \ldots, S^K\right\}$. Denote by $\mathbf{p}^{(n)}$ the n-th element of the action set $\mathcal{P}^{(b)}$. We write each vector $\mathbf{p}^{(n)}$ with $n \in \mathcal{I}$, as $\mathbf{p}^{(n)} = \left(\mathbf{p}_1^{(n)}, \ldots, \mathbf{p}_K^{(n)}\right)$, where for all $j \in \mathcal{K}$, $\mathbf{p}_j^{(n)} \in \mathcal{P}_j^{(b)}$. Consider that each action profile $\mathbf{p}^{(n)}$ is associated with a vertex v_n in a given non-directed graph G. Each vertex v_n is adjacent to the $K(S-1)$ vertices associated with the action profiles resulting when only one player deviates from the action profile $\mathbf{p}^{(n)}$, i.e., if two vertices v_n and v_m, with $(n, m) \in \mathcal{I}^2$ and $n \neq m$, are adjacent, then there exists one and only one $k \in \mathcal{K}$, such that

$$\forall j \in \mathcal{K} \setminus \{k\}, \quad \mathbf{p}_j^{(n)} = \mathbf{p}_j^{(m)}, \text{ and } \quad \mathbf{p}_k^{(n)} \neq \mathbf{p}_k^{(m)}.$$

More precisely, the graph G can be defined by the pair $G = (\mathcal{V}, \mathbf{A})$, where the set $\mathcal{V} = \{v_1, \ldots, v_{S^K}\}$ (nodes) represents the S^K possible actions profiles of the game and $\mathbf{A}$ (edges) is a symmetric matrix (adjacency matrix of G) with dimensions $S^K \times S^K$ and entries defined as follows $\forall (n, m) \in \mathcal{I}^2$ and $n \neq m$,

$$a_{n,m} = a_{m,n} = \begin{cases} 1 & \text{if } n \in \mathcal{V}_m \\ 0 & \text{otherwise} , \end{cases} \tag{3.18}$$

and $a_{n,n} = 0$ for all $n \in \mathcal{I}$, where the set $\mathcal{V}_n$ is the set of indices of the adjacent vertices of vertex v_n. In the following, we use the concept of distance between two vertices of the graph G. We define this concept using our notation:

Definition 3.4.1 (Shortest Path) *The distance (shortest path) between vertices v_n and v_m, with $(n, m) \in \mathcal{I}^2$ in a given non-directed graph $G = (\mathcal{V}, A)$, denoted by $d_{n,m}(G) \in \mathbb{N}$ is:*

$$d_{n,m}(G) = d_{m,n}(G) = \sum_{k=1}^{K} \mathbb{1}_{\left\{\mathbf{p}_k^{(n)} \neq \mathbf{p}_k^{(m)}\right\}}. \tag{3.19}$$

A realistic assumption is to consider that for any pair of action profiles $\mathbf{p}^{(n)}$ and $\mathbf{p}^{(m)}$, with $(n, m) \in \mathcal{I}^2$ and $n \neq m$, we have that $\phi(\mathbf{p}^{(n)}) \neq \phi(\mathbf{p}^{(m)})$ with probability one. This is because channel gains are random variables drawn from continuous probability distributions and thus, $\Pr\left(\phi(\mathbf{p}^{(n)}) = \phi(\mathbf{p}^{(m)}) \mid n \neq m\right) = 0$. Hence, following Def. 3.2.1, one can state that if the action profile $\mathbf{p}^{(n^*)}$, with $n^* \in \mathcal{I}$, is an NE of the game $\mathcal{G}_{(b)}$, then, it follows that

$$\forall m \in \mathcal{V}_{(n^*)}, \quad \phi(\mathbf{p}^{(n^*)}) > \phi(\mathbf{p}^{(m)}), \tag{3.20}$$

and vice versa with probability one. However, several action profiles might simultaneously satisfy the condition (3.20), which implies the non-uniqueness of the NE, as shown in the following proposition.

Proposition 3.4.2 *Let $\hat{K} \in \mathbb{N}$ be the highest even number inferior to K. Then, the game $\mathcal{G}_{(b)}$ has L pure NE strategy profiles, where,*

$$1 \leqslant L \leqslant 1 + \sum_{i \in \{2,4,\ldots,\hat{K}\}} \binom{K}{i} (S-1)\,. \tag{3.21}$$

Proof: From Prop. 3.4.1 it is ensured that $L \geqslant 1$. Then, assume that a given action profile $\mathbf{p}^{(n)}$ (vertex v_n) with $n \in \mathcal{I}$ is an NE. Given condition (3.20), it follows that none of the vertices in the set $\mathcal{V}_n$ is an NE. Nonetheless, if there exists another action profile $\mathbf{p}^{(m)}$, with $m \in \mathcal{I} \setminus \{n \cup \mathcal{V}_n\}$, which satisfies (3.20), then $\mathbf{p}^{(m)}$ can be also an NE. Thus, for the action profile $\mathbf{p}^{(m)}$, with $n \neq m$, to be an NE candidate, it must be (at least) at distance two of $\mathbf{p}^{(n)}$ and any other NE candidate, i.e., $d_{n,m}(G) = d_{m,n}(G) \in \{2, 4, \ldots, \hat{K}\}$. An action profile at distance $\ell \in \{2, 4, \ldots, \hat{K}\}$ from $\mathbf{p}^{(n)}$, is a vector where ℓ players have simultaneously deviated from $\mathbf{p}^{(n)}$. Hence, for each ℓ-tuple of players, there always exists $S-1$ action profiles at distance ℓ from $\mathbf{p}^{(n)}$ and at distance 2 from each other. Thus, aside from the initial NE action profile $\mathbf{p}^{(n)}$, there might exists at most

$$1 + \sum_{i \in \{2,4,\ldots,\hat{K}\}} \binom{K}{i} (S-1) \tag{3.22}$$

other NE candidates. This establishes an upper bound for L and completes the proof. □

The graph-theoretic interpretation of the channel selection problem allows one to see how big the number of pure NE can be. Note that this upper bound does not depend on the realizations of the channel gains whereas the number of NE generally depends on the latter. Interestingly, this upper bound (21) can be reached in the case $K = 2$ and $S = 2$, i.e., for certain channel realizations [86]. By contrast, in the case $K = 3$ and $S = 2$, the number of NE candidates is 4 (See Fig. 1 in [73]), however, as we shall see in Sec. 3.6, only 3 NE are observed. This is basically because in this case, there are some action profiles which are mutually exclusive from the set of NE. In any case, the upper-bound provided by Prop. 3.4.2 is tighter than the one in [73]. As in other games in the related literature (see e.g., the work by Scutari et al. [99]), expressing the number of NE for a given set of channel realizations is not easy. This is the reason why we will follow the approach of [99] by computing the probability of having a certain number of NE, which is done in Sec. 3.6.

3.4.3 Dynamics Arising from the Best Response

In order to fully identify the action profiles corresponding to an NE, we convert the non-directed graph G into an oriented graph $\hat{G}$ whose adjacency matrix is the non-symmetric square matrix $\hat{\mathbf{A}}$ whose entries are $\forall (i,j) \in \mathcal{I}^2$ and $i \neq j$,

$$\hat{a}_{i,j} = \begin{cases} 1 & \text{if} \quad i \in \mathcal{V}_j \text{ and } \phi\left(\mathbf{p}^{(j)}\right) > \phi\left(\mathbf{p}^{(i)}\right) \\ 0 & \text{otherwise}\,, \end{cases} \tag{3.23}$$

and $\hat{a}_{i,i} = 0$ for all $i \in \mathcal{I}$.
From the definition of the matrix $\hat{\mathbf{A}}$, we have that a necessary and sufficient condition for a vertex v_i to represent an NE action profile is to have a null out-degree in the oriented graph $\hat{G}$, i.e., there are no outgoing edges from the node v_i (sink vertex). Finally, one can conclude that determining the set of NE in the game $\mathcal{G}_{(b)}$ boils down to identifying all the sink vertices in the oriented graph $\hat{G}$. However, this task might be highly computationally demanding and requires complete information.

Assuming that the game is played repeatedly, it is known that the BRD converges to an NE in potential games [64]. The best response of transmitter $k \in \mathcal{K}$ in the game $\mathcal{G}_{(b)}$ is

$$\begin{aligned}\mathrm{BR}_k\left(\mathbf{p}_{-k}\right) \;=\; & \Big\{\mathbf{p}_k \in \mathcal{P}_k^{(b)} : \mathbf{p}_k = p_{k,\max}\mathbf{e}_{n_k^*} \text{ and } n_k^* = \\ & \arg\max_{s\in\mathcal{S}} \frac{B_S}{B}\log_2\left(1 + \frac{p_{k,\max}g_{k,s}}{\sigma_s^2 + \sum_{j\in\mathcal{K}\setminus\{k\}} p_{j,\max}g_{j,s}}\right)\Big\},\end{aligned}$$

and can be determined locally by each transmitter as described in Sec. 3.3.3. Additionally to the BRD (iterative and sequential, Def. 3.3.4), since the set of actions is discrete, other dynamics such as fictitious play [16] [86] can also be used.

3.4.4 Dynamics Arising from the Fictitious Play

The fictitious play can be described as follows. Assume that transmitters have complete and perfect information, i.e., they know the structure of the game $\mathcal{G}_{(b)}$ and observe at each time $t \in \mathbb{N}$ the PA vectors taken by all players. Each transmitter $k \in \mathcal{K}$ assumes that all its counterparts play independent and stationary (time-invariant) mixed strategies $\pi_j \in \triangle\left(\mathcal{P}_j^{(b)}\right)$, $\forall j \in \mathcal{K} \setminus \{k\}$. Under these conditions, player k is able to build an empirical probability distribution over each set $\mathcal{P}_j^{(b)}$, $\forall j \in \mathcal{K} \setminus \{k\}$. Let $f_{k,\mathbf{p}_k}(t) = \frac{1}{t}\sum_{s=1}^{t} \mathbb{1}_{\{\mathbf{p}_k(s)=\mathbf{p}_k\}}$ be the (empirical) probability with which players $j \in \mathcal{K} \setminus \{k\}$ observe that player k plays action $\mathbf{p}_k \in \mathcal{P}_k^{(b)}$. Hence, $\forall k \in \mathcal{K}$ and $\forall \mathbf{p}_k \in \mathcal{P}_k^{(b)}$, the following recursive expression holds,

$$f_{k,\mathbf{p}_k}(t+1) = f_{k,\mathbf{p}_k}(t) + \frac{1}{t+1}\left(\mathbb{1}_{\{\mathbf{p}_k(t)=\mathbf{p}_k\}} - f_{k,\mathbf{p}_k}(t)\right). \tag{3.24}$$

Let $\bar{f}_{k,\mathbf{p}_{-k}}(t) = \prod_{j\neq k} f_{j,\mathbf{p}_j}(t)$ be the probability with which player k observes the action profile $\mathbf{p}_{-k} \in \mathcal{P}_{-k}^{(b)}$ at time $t > 0$, for all $k \in \mathcal{K}$. Let the $\left|\mathcal{P}_{-k}^{(b)}\right|$ −dimensional vector $\mathbf{f}_k(t) = \left(\bar{f}_{k,\mathbf{p}_{-k}}\right)_{\forall \mathbf{p}_{-k}\in\mathcal{P}_{-k}^{(b)}} \in \triangle\left(\mathcal{P}_{-k}^{(b)}\right)$ be the empirical probability distribution over the set $\mathcal{P}_{-k}^{(b)}$ observed by player k. In the following, we refer to the vector $\mathbf{f}_k(t)$ as the *beliefs* of player k over the strategies of all its corresponding counterparts. Hence, based on its own beliefs $\mathbf{f}_k(t)$, each player k chooses its action at time t, $\mathbf{p}_k(t) = \mathbf{p}_k^{(n_k(t))}$, where $n_k(t)$ satisfies that:

$$n_k(t) \in \arg\max_{s\in\mathcal{S}} \bar{u}_k\left(\mathbf{e}_s, \mathbf{f}_k(t)\right), \tag{3.25}$$

where, for all $k \in \mathcal{K}$, $\bar{u}_k : \triangle\left(\mathcal{P}_1^{(b)}\right) \times \ldots \times \triangle\left(\mathcal{P}_K^{(b)}\right) \rightarrow \mathbb{R}$ is

$$\bar{u}_k(\pi) = \mathbb{E}_\pi\left[u_k(\mathbf{p}_k, \mathbf{p}_{-k})\right]. \tag{3.26}$$

From (3.24), it can be implied that playing FP, players become myopic, i.e., they build beliefs on the strategies being used by all the other players, and at each time $t > 0$, players choose the action that maximizes the instantaneous expected utility.

Convergence of the FP Dynamics

The game $\mathcal{G}_{(b)}$ is a potential game, and thus, it is said to have the fictitious play property (FPP) [63], i.e., for all $k \in \mathcal{K}$, and for all $\mathbf{p}_k \in \mathcal{P}_k^{(b)}$,

$$\lim_{t\rightarrow\infty} f_{k,\mathbf{p}_k}(t) = f^*_{k,\mathbf{p}_k}, \tag{3.27}$$

where, $\bar{f}^*_{k,\mathbf{p}_{-k}} = \prod_{j\in\mathcal{K}\setminus\{k\}} f^*_{j,\mathbf{p}_j}$, $\forall \mathbf{p}_{-k} \in \mathcal{P}_{-k}^{(b)}$, is a time-invariant probability distribution over the set $\mathcal{P}_{-k}^{(b)}$, which correspond to an NE in mixed strategies. An NE in mixed strategies, is defined as follows,

Definition 3.4.2 (Nash Equilibrium) *A mixed-strategy profile π^* is an NE of the game $\mathcal{G}_{(b)}$ if, for all players $k \in \mathcal{K}$ and $\forall \pi'_k \in \triangle\left(\mathcal{P}_k^{(b)}\right)$*

$$\bar{u}_k(\pi^*_k, \pi^*_{-k}) \geqslant \bar{u}_k(\mathbf{pi}'_k, \pi^*_{-k}). \tag{3.28}$$

Here, for all $k \in \mathcal{K}$, $\pi_k = \left(\pi_{k,\mathbf{p}_k^{(1)}}, \ldots, \pi_{k,\mathbf{p}_k^{(S)}}\right) \in \triangle\left(\mathcal{P}_k^{(b)}\right)$, where $\forall (k,s) \in \mathcal{K} \times \mathcal{S}$, $\pi_{k,\mathbf{p}_k^{(s)}}$ represents the probability that player k uses channel s.
Following this reasoning, we write the following proposition

Proposition 3.4.3 (Convergence of FP in CS) *The fictitious play converges empirically to the set of Nash equilibrium in the game $\mathcal{G}_{(b)}$.*

Practical Limitations of Fictitious Play

As presented in its original version [16], the FP requires complete and perfect information. This is the same as stating that each transmitter, at each time $t > 0$, is aware of the number of active transmitters in the network, their set of actions, their utility function and moreover, it is able to observe the action played by each one of all the other transmitters. Clearly, this assumption is not practically appealing since it would require a massive signaling between transmitters, which reduces the spectral efficiency of the whole network. Additionally, as we shall see, in the high SNR regime, the CS problem has the same structure of a potential coordination game [119]. In this kind of games, the set of probability distributions $\mathbf{f}_k$, $\forall k \in \mathcal{K}$, converges but not necessarily the actions, i.e., fictitious play might converge to a strictly mixed strategy profile. When FP converges to a mixed strategy, it is possible that players cycle around a subset of action profiles, which might lead to an expected utility which is worse that the worst expected utility at the NE in pure and mixed strategies. In the following section, we present a simple study case where it is easy to evidence this cycling effect.

3.5 Analysis of Special Cases

In this section, we study two special cases of relevant interest to understand previous conclusions. First, the games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ are studied assuming that there exist only $K = 2$ transmitters and $S = 2$ available channels. In particular, we analyze the set of NE action profiles of both games and compare the network spectral efficiency (NSE) $U : \mathcal{P}^{(b)} \to \mathbb{R}$, that is obtained by playing both games. Here,

$$U(\mathbf{p}_1, \ldots, \mathbf{p}_K) = \sum_{k=1}^{K} u_k(\mathbf{p}_1, \ldots, \mathbf{p}_K) \text{ [bps/Hz]}. \tag{3.29}$$

From this analysis, it is concluded that from a network performance point of view, limiting the transmitters to use a unique channel brings a better result in terms of network spectral efficiency (3.29). Second, we consider the case of a large number of transmitters and channels assuming a finite ratio of transmitters per channel. This study leads to conclude that, the fraction of players using a given channel depends mainly on the bandwidth of each channel.

3.5.1 The 2-Transmitter 2-Channel Case

Consider the games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ with $K = 2$ and $S = 2$. Assume also that $\forall k \in \mathcal{K}$, $p_{k,\max} = p_{\max}$ and $\forall s \in \mathcal{S}$, $\sigma_s^2 = \sigma^2$ and $B_s = \frac{B}{S}$. Denote by $\mathrm{SNR} = \frac{p_{\max}}{\sigma^2}$ the average signal to noise ratio (SNR) of each active communication.

The Power Allocation Game

Let us denote by $\mathbf{p}^\dagger = \left(\mathbf{p}_1^\dagger, \mathbf{p}_2^\dagger\right)$ the NE of the game $\mathcal{G}_{(a)}$. Then, following Def. 3.2.1, one can write the following set of inclusions,

$$\forall k \in \mathcal{K}, \quad \mathbf{p}_k^\dagger \in \mathrm{BR}_k\left(\mathbf{p}_{-k}^\dagger\right). \tag{3.30}$$

Note that, for all $k \in \mathcal{K}$ and for all $\mathbf{p}_{-k} \in \mathcal{P}^{(a)}$, the set $\mathrm{BR}_k\left(\mathbf{p}_{-k}\right)$ is a singleton (Def. 3.3.3) and thus, (3.30) represents a system of equations. By solving the resulting system of equations (3.30) for a given realization of the channels $\{g_{i,j}\}_{\forall (i,j) \in \mathcal{K} \times \mathcal{P}}$, one can determine the NE of the game $\mathcal{G}_{(a)}$. We present such a solution in the following proposition.

Proposition 3.5.1 (Nash Equilibrium in $\mathcal{G}_{(a)}$) *Let the action profile $\mathbf{p}^\dagger = \left(\mathbf{p}_1^\dagger, \mathbf{p}_2^\dagger\right) \in \mathcal{P}^{(a)}$, with $\mathbf{p}_1^\dagger = \left(p_{11}^\dagger, p_{\max} - p_{11}^\dagger\right)$ and $\mathbf{p}_2^\dagger = \left(p_{\max} - p_{22}^\dagger, p_{22}^\dagger\right)$ be an NE action profile of the game $\mathcal{G}_{(a)}$. Then, with probability one, $\mathbf{p}^\dagger$ is the unique NE and it can be written as follows:*

- *Equilibrium 1: if $\mathbf{g} \in \mathcal{B}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}, \frac{g_{21}}{g_{22}} \leqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}\}$, then, $p_{11}^\dagger = p_{\max}$ and $p_{22}^\dagger = p_{\max}$.*
- *Equilibrium 2: if $\mathbf{g} \in \mathcal{B}_2 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right), \frac{g_{21}}{g_{22}} \geqslant 1 + \mathrm{SNR}\left(g_{11} + g_{21}\right)\}$, then, $p_{11}^\dagger = p_{\max}$ and $p_{22}^\dagger = 0$.*

- *Equilibrium* 3*: if* $\mathbf{g} \in \mathcal{B}_3 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})}, \frac{g_{21}}{g_{22}} \leqslant \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})}\}$ *then,* $p_{11}^{\dagger} = 0$ *and* $p_{22}^{\dagger} = p_{\max}$.
- *Equilibrium* 4*: if* $\mathbf{g} \in \mathcal{B}_4 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}, \frac{g_{21}}{g_{22}} \geqslant \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}\}$*, then,* $p_{11}^{\dagger} = 0$ *and* $p_{22}^{\dagger} = 0$.
- *Equilibrium* 5*: if* $\mathbf{g} \in \mathcal{B}_5\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{g_{21}}{g_{22}}, \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}} < \frac{g_{21}}{g_{22}} < 1 + \mathrm{SNR}\left(g_{11}+g_{21}\right)\}$*, then,* $p_{11}^{\dagger} = p_{\max}$ *and* $p_{22}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{22}} + \frac{\sigma^2 + g_{11}p_{\max}}{g_{21}}\right)$.
- *Equilibrium* 6*: if* $\mathbf{g} \in \mathcal{B}_6\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{g_{21}}{g_{22}}, \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})} < \frac{g_{11}}{g_{12}} < \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}\}$*, then,* $p_{11}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{11}} + \frac{\sigma^2 + p_{\max}g_{22}}{g_{12}}\right)$ *and* $p_{22}^{\dagger} = p_{\max}$.
- *Equilibrium* 7*: if* $\mathbf{g} \in \mathcal{B}_7 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{22}}, \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}} < \frac{g_{11}}{g_{12}} < 1 + \mathrm{SNR}\left(g_{11}+g_{21}\right)\}$*, then,* $p_{11}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + p_{\max}g_{21}}{g_{11}} + \frac{\sigma^2}{g_{12}}\right)$ *and* $p_{22}^{\dagger} = 0$.
- *Equilibrium* 8*: if* $\mathbf{g} \in \mathcal{B}_8\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{22}}, \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})} < \frac{g_{21}}{g_{22}} < \frac{1+\mathrm{SNR}g_{21}}{1+\mathrm{SNR}g_{12}}\}$*, then,* $p_{11}^{\dagger} = 0$ *and* $p_{22}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{22}} + \frac{\sigma^2}{g_{21}}\right)$.

Proof: See Appendix D □

In Fig. 3.1 we plot the different types of NE of the game $\mathcal{G}_{(a)}$ as a function of the channel ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$. The uniqueness of the NE is not ensured under certain conditions as we show in the following proposition. In fact, under those conditions infinitely many NE can be observed, however, such conditions are observed with zero probability.

Proposition 3.5.2 *Assume that the set of channels* $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$ *satisfies the following conditions*

$$\frac{1}{1+\frac{p_{\max}}{\sigma^2}(g_{12}+g_{22})} < \frac{g_{11}}{g_{12}} = \frac{g_{21}}{g_{22}} < 1 + \frac{p_{\max}}{\sigma^2}(g_{11}+g_{21}), \tag{3.31}$$

Then, any PA vector $\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathcal{P}^{(a)}$*, such that*

$$p_{11} = \frac{1}{2}\left(p_{\max}\left(1-\alpha\right) + \sigma^2\left(\frac{1}{g_{12}} - \frac{1}{g_{11}}\right)\right) + \alpha p_{22}$$

with $\alpha \triangleq \frac{g_{11}}{g_{21}} = \frac{g_{12}}{g_{22}}$*, is an NE action profile of the game* $\mathcal{G}_{(a)}$.

The proof of Prop. 3.5.2 is the first part of the proof of Prop. 3.5.1.

In the next subsection, we perform the same analysis presented above for the game $\mathcal{G}_{(b)}$.

The Channel Selection Game

When $K = 2$ and $S = 2$, the game $\mathcal{G}_{(b)}$ has four possible outcomes, i.e., $\left|\mathcal{P}^{(b)}\right| = 4$. We detail such outcomes and its respective potential in Fig. 3.2.

Following Def. 3.2.1, each of those outcomes can be an NE depending on the channel realizations $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$, as shown in the following proposition.

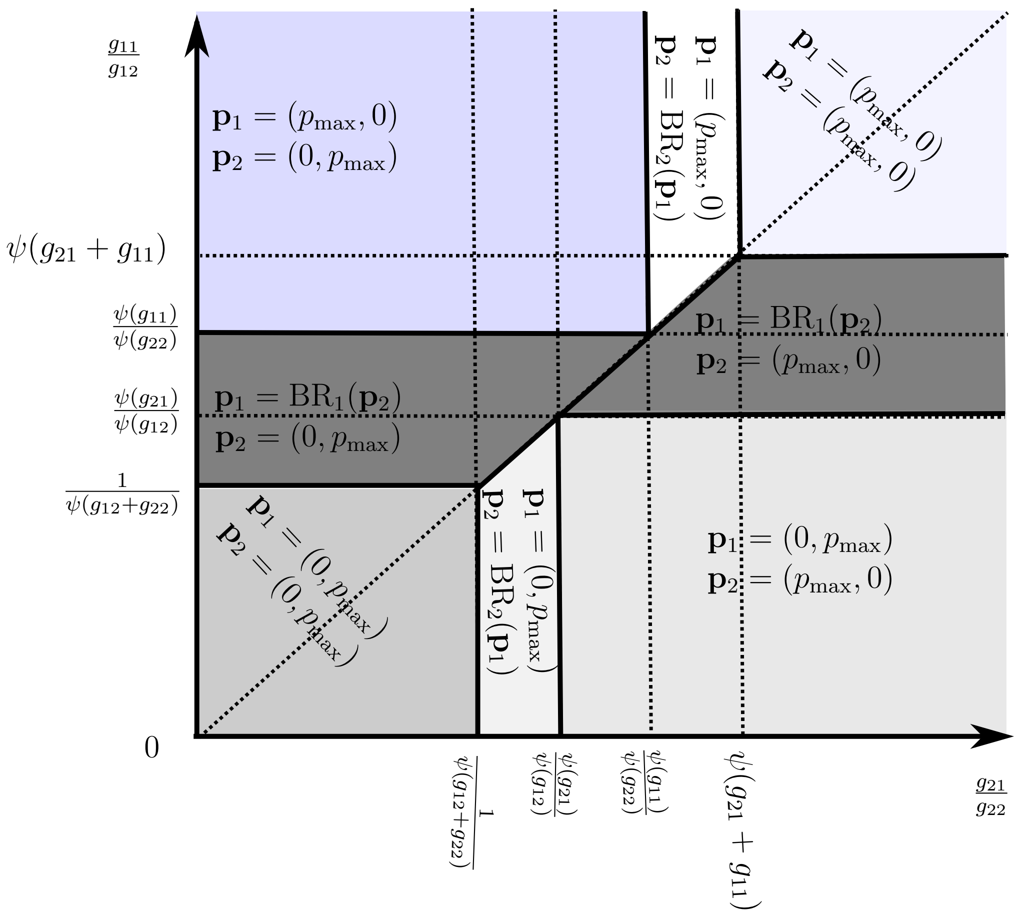

Figure 3.1: Nash equilibrium action profiles as a function of the channel ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$ for the two-player-two-channel game $\mathcal{G}_{(a)}$. The function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as follows: $\psi(x) = 1 + \mathrm{SNR}\, x$. The best response function $\mathrm{BR}_k(\mathbf{p}_{-k})$, for all $k \in \mathcal{K}$, is defined by (3.14). Here, it has been arbitrarily assumed that $\frac{\psi(g_{21})}{\psi(g_{12})} < \frac{\psi(g_{11})}{\psi(g_{22})}$.

$Tx_1 \backslash Tx_2$	$\mathbf{p}_2=(p_{\max},0)$	$\mathbf{p}_2=(0,p_{\max})$
$\mathbf{p}_1=(p_{\max},0)$	$\frac{1}{2}\log_2\left(\sigma^2+p_{\max}(g_{11}+g_{21})\right)+\frac{1}{2}\log_2\left(\sigma^2\right)$	$\frac{1}{2}\log_2\left(\sigma^2+p_{\max}g_{11}\right)+\frac{1}{2}\log_2\left(\sigma^2+p_{\max}g_{22}\right)$
$\mathbf{p}_1=(0,p_{\max})$	$\frac{1}{2}\log_2\left(\sigma^2+p_{\max}g_{12}\right)+\frac{1}{2}\log_2\left(\sigma^2+p_{\max}g_{21}\right)$	$\frac{1}{2}\log_2\left(\sigma^2+p_{\max}(g_{12}+g_{22})\right)+\frac{1}{2}\log_2\left(\sigma^2\right)$

Figure 3.2: Potential function ϕ of the game $\mathcal{G}_{(b)}$, with $K = 2$ and $S = 2$. Player 1 chooses rows and player 2 chooses columns.

Proposition 3.5.3 (Nash Equilibria in $\mathcal{G}_{(b)}$) *Let the PA vector $\mathbf{p}^* = (\mathbf{p}_1^*, \mathbf{p}_2^*) \in \mathcal{P}^{(b)}$ be one NE in the game $\mathcal{G}_{(b)}$. Then, depending on the channel gains $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$, the NE $\mathbf{p}^*$ can be written as follows:*

- *Equilibrium 1: when $\mathbf{g} \in \mathcal{A}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{1}{1+\mathrm{SNR}g_{22}}$ and $\frac{g_{21}}{g_{22}} \leqslant 1 + \mathrm{SNR}g_{11}\}$, then, $\mathbf{p}_1^* = (p_{\max}, 0)$ and $\mathbf{p}_2^* = (0, p_{\max})$.*
- *Equilibrium 2: When $\mathbf{g} \in \mathcal{A}_2 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 + \mathrm{SNR}g_{21}$ and $\frac{g_{21}}{g_{22}} \geqslant 1 + \mathrm{SNR}g_{11}\ \}$, then, $\mathbf{p}_1^* = (p_{\max}, 0)$ and $\mathbf{p}_2^* = (p_{\max}, 0)$.*
- *Equilibrium 3: when $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{A}_3\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{1}{1+\mathrm{SNR}g_{22}}$ and $\frac{g_{21}}{g_{22}} \leqslant \frac{1}{1+\mathrm{SNR}g_{12}}\}$, then, $\mathbf{p}_1^* = (0, p_{\max})$ and $\mathbf{p}_2^* = (0, p_{\max})$.*
- *Equilibrium 4: when $\mathbf{g} \in \mathcal{A}_4 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant 1 + \mathrm{SNR}g_{12}$ and $\frac{g_{21}}{g_{22}} \geqslant \frac{1}{1+\mathrm{SNR}g_{12}}\ \}$, then, $\mathbf{p}_1^* = (0, p_{\max})$ and $\mathbf{p}_2^* = (p_{\max}, 0)$.*

Proof: The proof follows immediately from Def. 3.2.1 and Fig. 3.2. □

In Fig. 3.3, we plot the different types of NE action profiles as a function of the channel ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$. Note how the action profiles $\mathbf{p}^* = (p_{\max}, 0, 0, p_{\max})$ and $\mathbf{p}^+ = (0, p_{\max}, p_{\max}, 0)$ are both NE, when the channel realizations satisfy that $\mathbf{g} \in \mathcal{A}_5 = \mathcal{A}_1 \cap \mathcal{A}_4$, i.e., $\mathcal{A}_5 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{1}{1+\mathrm{SNR}g_{22}} \leqslant \frac{g_{11}}{g_{12}} \leqslant 1+\mathrm{SNR}g_{21}$ and $\frac{1}{1+\mathrm{SNR}g_{12}} \leqslant \frac{g_{21}}{g_{22}} \leqslant 1+\mathrm{SNR}g_{11}\}$. This confirms the fact that several NE might exists in the game $\mathcal{G}_{(b)}$ depending on the exact channel realization, as stated in Prop. 3.4.2. Moreover, one can also observe that there might exist an NE action profile which is not a global maximizer of the potential function (3.10) [113] (e.g., $\phi(\mathbf{p}^*) < \phi\left(\mathbf{p}_2^+\right)$).

Now, following the result in [117], it can be implied that when there exist two NE in pure strategies, there exists a third NE in mixed strategies. When, there exists a unique NE in pure strategies, the NE in mixed strategies coincides with the NE in pure strategies. We summarize this observation in the following proposition.

Proposition 3.5.4 (NE in Mixed Strategies) *Let $\pi_k^* \in \triangle\left(\mathcal{P}_k^{(b)}\right)$ be a mixed strategy of player k, $\forall k \in \mathcal{K}$. Then, $\pi^* = (\pi_1^*, \ldots, \pi_K^*)$ is an NE in strict mixed strategies of the game $\mathcal{G}_{(b)}$, if and only if, the channel realizations $\{g_{i,j}\}_{\forall(i,j)\in\mathcal{K}\times\mathcal{P}}$*

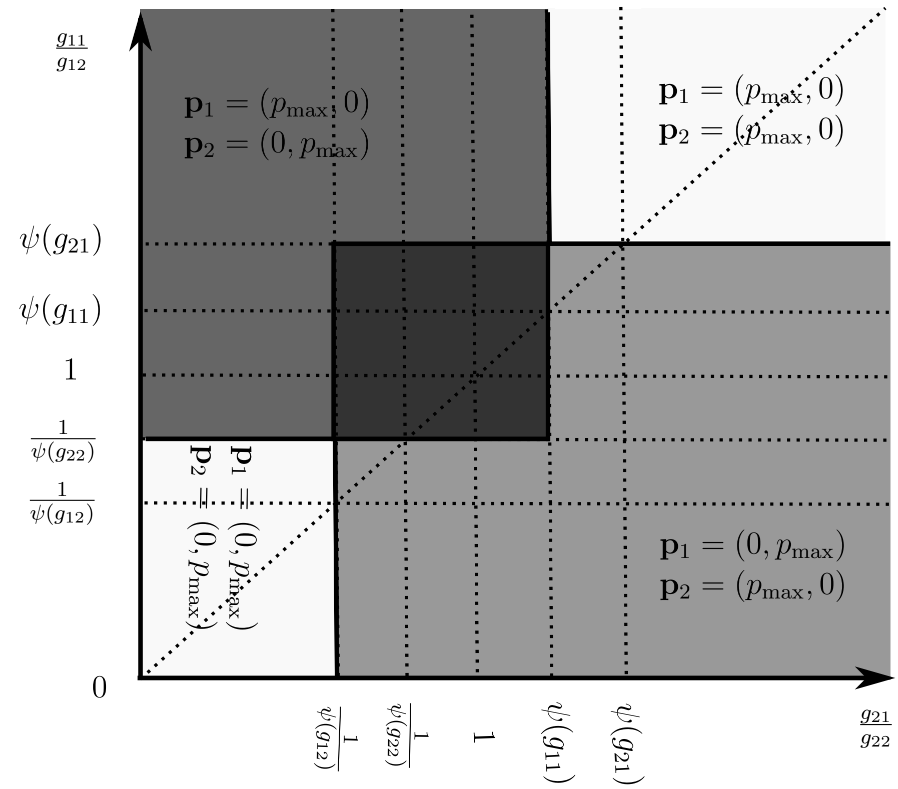

Figure 3.3: Nash equilibrium action profiles as a function of the ratios $\frac{g_{11}}{g_{12}}$ and $\frac{g_{21}}{g_{22}}$ of the channel realizations for the two-player-two-channel game $\mathcal{G}_{(b)}$. The function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ is defined as follows: $\psi(x) = 1 + \mathrm{SNR}\, x$. Here, it has been arbitrarily assumed that $\psi(g_{11}) < \psi(g_{21})$.

satisfy that $\mathbf{g} \in \mathcal{A}_1 \cap \mathcal{A}_4$ *and,*

$$\pi^*_{1,\mathbf{p}^{(1)}} = \frac{\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) + \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}, \tag{3.32}$$

$$\pi^*_{1,\mathbf{p}^{(2)}} = \frac{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right)}{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) + \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}, \tag{3.33}$$

$$\pi^*_{2,\mathbf{p}^{(1)}} = \frac{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) + \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}, \tag{3.34}$$

$$\pi^*_{2,\mathbf{p}^{(2)}} = \frac{\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right)}{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right) + \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}. \tag{3.35}$$

where, $\mathbf{p}^{(1)} = (0, p_{\max})$ *and* $\mathbf{p}^{(2)} = (p_{\max}, 0)$

In the sequel, the performance achieved by the transmitters at the equilibrium in both games are compared.

A note on the Convergence of Fictitious Play

In the case the NE is unique in the CS game $\mathcal{G}_{(b)}$, the FP converges to the unique NE in pure strategies (Prop. 3.4.3). Nonetheless, when several NE simultaneously exist, the FP converges to the NE either in pure strategies or mixed strategies. In the following, we show a case of convergence in mixed strategies using the FP. Assume that both players starts the game with the initial beliefs

$$\mathbf{f}_j(t_0) = \left(f_{j,\mathbf{p}^{(1)}}(t_0), f_{j,\mathbf{p}^{(2)}}(t_0)\right),$$

such that $f_{j,\mathbf{p}^{(1)}}(t_0) = \frac{\xi_j}{1+\xi_j}$ and $f_{j,\mathbf{p}^{(2)}}(t_0) = \frac{1}{1+\xi_j}$, with $0 < \xi_j < 1$, for all $j \in \mathcal{K}$. Hence, based on these beliefs, both players coincide choosing the action $\mathbf{p}^{(1)}$ at $t = t_0$. Following (3.24), it yields, $\forall k \in \mathcal{K}$, and $\forall n \in \{1,\ldots,\infty\}$,

$$\left\{\begin{array}{lcl} f_{k,\mathbf{p}^{(1)}}(t_0{+}2n{-}1) & = & \frac{1}{2n-1}\left(\frac{n\xi_k+(n-1)}{1+\xi_k}\right) \\ f_{k,\mathbf{p}^{(2)}}(t_0{+}2n{-}1) & = & \frac{1}{2n-1}\left(\frac{(n-1)\xi_k+n}{1+\xi_k}\right) \\ f_{k,\mathbf{p}^{(1)}}(t_0{+}2n) & = & \frac{1}{2n}\left(\frac{(n+1)\xi_k+n}{1+\xi_k}\right) \\ f_{k,\mathbf{p}^{(2)}}(t_0{+}2n) & = & \frac{1}{2n}\left(\frac{(n-1)\xi_k+n}{1+\xi_k}\right) \end{array}\right. . \tag{3.36}$$

Here, as long as the following condition holds $\forall k \in \mathcal{K}$ and a given $n \in \{1,\ldots,\infty\}$,

$$\frac{n(\xi_k+1)-1}{n(\xi_k+1)-\xi_k} \leqslant \frac{\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right)-\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)}{\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right)-\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right)} \leqslant \frac{n(\xi_k+1)+\xi_k}{n(\xi_k+1)-\xi_k}, \tag{3.37}$$

then, the following outcomes are observed,

$$\mathbf{p}_k(2n-1) = \mathbf{p}^{(1)} \quad \text{and} \quad \mathbf{p}_k(2n) = \mathbf{p}^{(2)}.$$

This implies that transmitters will cycle around the outcomes $\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right)$ and $\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)$. Note that if

$$\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(1)}\right)-\phi\left(\mathbf{p}^{(2)},\mathbf{p}^{(2)}\right)=\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(2)}\right)-\phi\left(\mathbf{p}^{(1)},\mathbf{p}^{(1)}\right), \tag{3.38}$$

then, the beliefs of each player converge to $\pi_{k,\mathbf{p}^{(s)}} = \frac{1}{2}$, for all $(k,s) \in \mathcal{K} \times \mathcal{S}$ and players perpetually iterate between actions $\left(\mathbf{p}^{(1)}, \mathbf{p}^{(1)}\right)$ and $\left(\mathbf{p}^{(2)}, \mathbf{p}^{(2)}\right)$. Here, even though $\pi_k = \left(\frac{1}{2}, \frac{1}{2}\right)$, for all $k \in \mathcal{K}$, is an NE in mixed strategies according to Prop. 3.5.4, the achieved expected utility can be worse than the worst expected utility at NE in pure and mixed strategies. This can be explained by the fact that the pure strategies corresponding to the NE, i.e., $\mathbf{p}^{\dagger} = \left(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}\right)$ and $\mathbf{p}^{\dagger\dagger} = \left(\mathbf{p}^{(2)}, \mathbf{p}^{(1)}\right)$, are never played. Hence, if the channel realizations are those such that sharing the same channel is always worse than using orthogonal channels, i.e., $\phi\left(\mathbf{p}^{(2)}, \mathbf{p}^{(1)}\right) >> \phi\left(\mathbf{p}^{(2)}, \mathbf{p}^{(2)}\right)$ and $\phi\left(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}\right) >> \phi\left(\mathbf{p}^{(1)}, \mathbf{p}^{(1)}\right)$, then, a worse utility than the worst NE either in pure or mixed strategies is observed.
Interestingly, if the differences $\phi\left(\mathbf{p}^{(2)}, \mathbf{p}^{(1)}\right) - \phi\left(\mathbf{p}^{(2)}, \mathbf{p}^{(2)}\right)$ and $\phi\left(\mathbf{p}^{(1)}, \mathbf{p}^{(2)}\right) - \phi\left(\mathbf{p}^{(1)}, \mathbf{p}^{(1)}\right)$ are sufficiently close, then, a large number n in (3.37) is required for the FP to quit the cycle mentioned above. This implies that a long time is required for players to play the four actions profiles and thus, obtain the expected utility corresponding to the NE in mixed strategies. Here, as long as $\phi(\mathbf{p}^{(2)},\mathbf{p}^{(1)}) - \phi(\mathbf{p}^{(2)},\mathbf{p}^{(2)}) \neq \phi(\mathbf{p}^{(1)},\mathbf{p}^{(2)}) - \phi(\mathbf{p}^{(1)},\mathbf{p}^{(1)})$, there always exists an $n_0 < \infty$, such that $\forall n > n_0$, condition (3.38) does not hold, and thus, the cycling effect is not longer observed.

3.5.2 A Braess Type Paradox

In the game $\mathcal{G}_{(b)}$, the set of actions for player k is a subset of its set of actions in the game $\mathcal{G}_{(a)}$, i.e., $\forall k \in \mathcal{K}$, $\mathcal{P}_k^{(b)} \subseteq \mathcal{P}_k^{(a)}$. Thus one might think that having a larger set of actions might lead to a better global performance, for instance, a higher NSE. In this section, we show that contrary to intuition, reducing the set of actions of each player might lead to a better global performance. This effect (often associated with a Braess paradox [15]) has been reported in the parallel interference channel in [90] and in the parallel MAC for the case of successive interference cancellation (SIC) in [75]. In the following, we do not impose any condition on the channel gains and compare the NSE obtained by either playing the game $\mathcal{G}_{(a)}$ or $\mathcal{G}_{(b)}$ at the NE to study this counter intuitive result.
Let us denote by $\mathbf{p}_k^{(\dagger,n)}$, the unique NE action profile of game $\mathcal{G}_{(a)}$, when the vector $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_n$, for all $n \in \{1\ldots, 8\}$. Let us also denote by $\mathbf{p}^{(*,n)}$ one of the NE action profiles of game $\mathcal{G}_{(b)}$ when $(g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{A}_n$, for all $n \in \{1, \ldots 4\}$. The sets $\mathcal{A}_n$ and $\mathcal{B}_n$ are defined in Prop. 3.5.1 and 3.5.3. Then, for a finite SNR level, SNR > 0, one can observe that $\forall n \in \{1, \ldots, 4\}$, $\mathcal{A}_n \cap \mathcal{B}_n = \mathcal{B}_n$ and $\forall \mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_n$, the following equality always holds $\mathbf{p}_k^{(\dagger,n)} = \mathbf{p}_k^{(*,n)}$, which implies the same network performance. However, when the NE of both games are different, one can not easily compare the utilities achieved by each player since they depend on the exact channel realizations. Fortunately, the analysis largely simplifies by considering either a low SNR regime or a high SNR regime and more general conclusions can be stated. The performance comparison between games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ for the low SNR regime is presented in the following proposition.

Proposition 3.5.5 *In the low SNR regime, both games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ possess a* *unique NE, denoted by* $\mathbf{p}^*$. *Here, for all* $k \in \mathcal{K}$ *and* $n_k \in \mathcal{S}$,

$$\begin{aligned} p^*_{k,n_k} &= p_{\max} \mathbb{1}_{\left\{n_k = \arg\max\limits_{\ell \in \mathcal{S}} g_{k,\ell}\right\}} && (3.39)\\ p^*_{k,-n_k} &= p_{\max} - p_{k,n_k}. && (3.40) \end{aligned}$$

Proof: See App. F □

Similarly, the performance comparison between games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ for the high SNR regime is presented in the following proposition.

Proposition 3.5.6 (Existence of a Braess paradox) *In the high SNR regime, the game* $\mathcal{G}_{(a)}$ *has a unique pure NE (denoted by* $\mathbf{p}^\dagger$*) and the game* $\mathcal{G}_{(b)}$ *has two pure NE (denoted by* $\mathbf{p}^{(*,1)}$ *and* $\mathbf{p}^{(*,4)}$*). Then, at least for one* $n \in \{1,4\}$, $\exists \mathrm{SNR}_0 > 0$, *such that* $\forall \mathrm{SNR} \geqslant \mathrm{SNR}_0$,

$$\sum_{k=1}^{2} u_k\left(\mathbf{p}^{(*,n)}\right) - \sum_{k=1}^{2} u_k(\mathbf{p}^\dagger) \geqslant \delta, \qquad (3.41)$$

and $\delta \geqslant 0$.

For the proof see App. G. Note that from Prop. 3.5.5 and Prop. 3.5.6, it can be concluded that at low SNR both games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$ induce the same network spectral efficiency. On the contrary, the game $\mathcal{G}_{(b)}$ always induce a higher or equal network spectral efficiency than the game $\mathcal{G}_{(a)}$ in the high SNR regime. This counter-intuitive result implies a Braess type paradox, since $\mathcal{P}^{(b)} \subset \mathcal{P}^{(a)}$.

3.5.3 The Case of Large Systems

In this section, we exclusively deal with the game $\mathcal{G}_{(b)}$ for the case of large networks, i.e., networks with a large number of transmitters and available channels. Typically, under this condition, the dominant parameter is the number of transmitters per channel $\eta = \frac{K}{S}$. As we shall see, contrary to the case of small number of transmitters and channels analyzed in the preceding section, in the case of large networks, each player becomes indifferent to the action adopted by each of the other players. Here, each player is rather concerned with the fractions of players simultaneously playing the same action. Hence, one of the interesting issues to be solved is the determination of the repartition of the users between the different channels at the NE. Like in routing problems, transmitters are attracted by the best channels (the highest channel gains and largest bands) but the corresponding channels can turn out to be "too crowded". As a first step towards identifying the fractions of transmitters per channel at the NE, we first re-write the potential (3.10) as a function of the vector $\mathbf{x}(\mathbf{p}) = (x_1(\mathbf{p}), \ldots, x_S(\mathbf{p}))$, where $x_s(\mathbf{p})$, with $s \in \mathcal{S}$, denotes the fraction of players transmitting over channel s given the action profile $\mathbf{p} \in \mathcal{P}^{(b)}$. Hence,

$$\begin{aligned} \forall s \in \mathcal{S}_k, \quad x_s(\mathbf{p}) &= \frac{|\mathcal{K}_s(\mathbf{p})|}{K} \\ \sum_{i=1}^{S} x_i(\mathbf{p}) &= 1, \end{aligned} \qquad (3.42)$$

where $\mathcal{K}_s(\mathbf{p}) \subseteq \mathcal{K}$ is the set of players using channel s given the action profile $\mathbf{p} \in \mathcal{P}^{(b)}$, i.e., $\mathcal{K}_s(\mathbf{p}) = \{k \in \mathcal{K} : p_{k,s} \neq 0\}$. Let $b_s = \frac{B_s}{B}$ denote the fraction of bandwidth associated with channel s, such that $\sum_{s=1}^{S} b_s = 1$ and let $\bar{B} = \frac{B}{S}$ denote the mean bandwidth per channel. Then, one can write the potential as follows

$$
\begin{aligned}
\phi(\mathbf{p}) &= \sum_{s=1}^{S} b_s \log_2\left(N_0 B_s + p_{\max} \sum_{k\in\mathcal{K}_s(\mathbf{p})} g_{k,s}\right)\\
&= \sum_{s=1}^{S} b_s \log_2\left(\frac{N_0 \bar{B}}{\eta} b_s + x_s(\mathbf{p})\, p_{\max}\left(\frac{1}{|\mathcal{K}_s(\mathbf{p})|}\sum_{k\in\mathcal{K}_s(\mathbf{p})} g_{k,s}\right)\right)\\
&\quad + \sum_{s=1}^{S} b_s \log_2(K).
\end{aligned} \tag{3.43}
$$

Note that under the assumption of large number of transmitters, the following approximation holds, independently of the action profile $\mathbf{p}$,

$$
\forall s \in \mathcal{S},\ \frac{1}{|\mathcal{K}_s(\mathbf{p})|}\sum_{k\in\mathcal{K}_s(\mathbf{p})} g_{k,s} \approx \int_0^{\infty} \lambda \mathrm{d}F_{g_s}(\lambda) = \Omega_s,
$$

where F_{g_s} is the cumulative probability function associated with the channel gains over dimension s. Hence, independently of the exact action profile $\mathbf{p} \in \mathcal{P}^{(b)}$ adopted by the players,

$$
\begin{aligned}
\phi\left(\mathbf{p}\right) \approx \tilde{\phi}\left(\mathbf{x}(\mathbf{p})\right) = & \sum_{s=1}^{S} \log_2\left(\frac{N_0 \bar{B}}{\eta} b_s + x_s(\mathbf{p})\, p_{\max}\, \Omega_s\right)\\
& + \sum_{s=1}^{S} b_s \log_2(K).
\end{aligned} \tag{3.44}
$$

Finally, finding an NE in the non-atomic extension of the game $\mathcal{G}_{(b)}$, i.e., the set of fractions $x_1, \ldots, x_S$, which maximizes the potential of the game, boils down to solving the following OP (Def. 3.2.2),

$$
\begin{cases}
\max\limits_{\mathbf{x}\in\mathbb{R}_+^S} & \sum_{s=1}^{S} b_s \log_2\left(\frac{N_0 \bar{B}}{\eta}\, b_s + x_s p_{\max} \Omega_s\right),\\
\text{s.t.} & \sum_{i=1}^{S} x_i = 1 \ \text{ and } \ \forall i \in \mathcal{S},\, x_i \geqslant 0,
\end{cases} . \tag{3.45}
$$

The optimization problem (3.45) has a unique solution of the form,

$$
\forall s \in \mathcal{S}, \quad x_s = b_s \left(\frac{1}{\beta} - \frac{N_0 \bar{B}}{\eta p_{\max} \Omega_s}\right)^{+}, \tag{3.46}
$$

where β_k is Lagrangian multiplier to satisfy the optimization constraints. Interestingly, in the case when $\forall s \in \mathcal{S}$, $F_{g_s}(\lambda) = F_g(\lambda)$, $(\forall s \in \mathcal{S},\ \Omega_s = \Omega)$ it holds that,

$$
\forall s \in \mathcal{S}, \quad x_s = \frac{B_s}{B}. \tag{3.47}
$$

In this case, the solution does not depend on the fraction of players per channel $\eta = \frac{K}{S}$, but only in the distribution of the bandwidth among all the channels. For instance, assuming $\forall (k,s) \in \mathcal{K} \times \mathcal{S}$, $g_{k,s} = 1$ and $\forall s \in \mathcal{S}$, $B_s = \frac{B}{S}$, the same result presented in [8] is obtained.

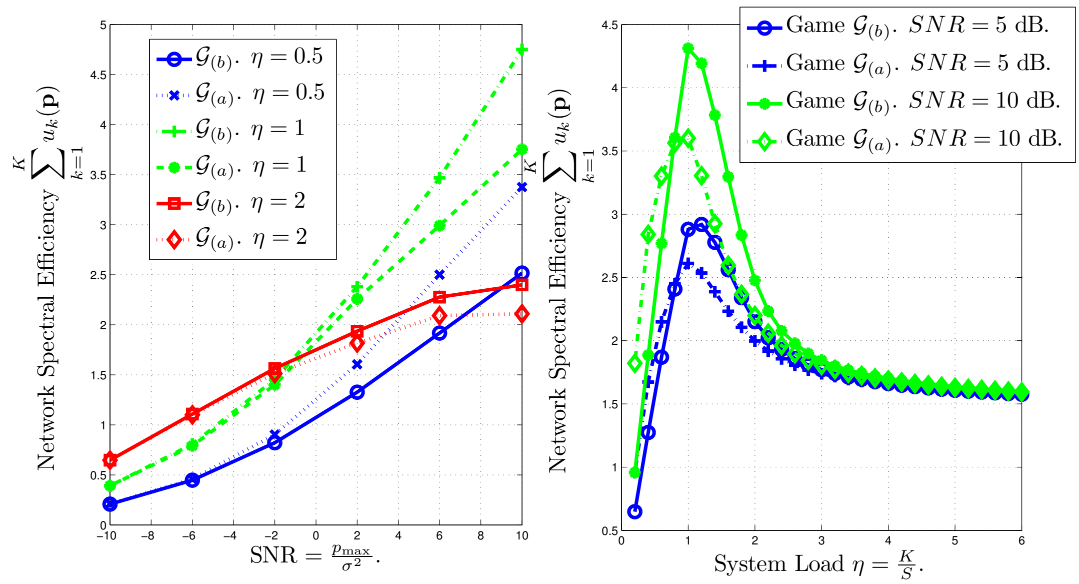

Figure 3.4: (a) Network spectral efficiency as a function of the system load $\eta = \frac{K}{S}$ for different SNR $= \frac{p_{\max}}{\sigma^2}$ levels in dBs. (b) Network spectral efficiency as a function of the SNR $= \frac{p_{\max}}{\sigma^2}$ in dBs. for the case of $\eta = \frac{K}{S} \in \{\frac{1}{2}, 1, \frac{3}{2}\}$, with $K = 10$.

3.6 Numerical Examples

In the previous sections, a mathematical argument has been provided to show that at the low and high SNR regime, transmitting over a single channel at the maximum power yields a higher NSE, at least for the case of $K = 2$ transmitters and $S = 2$ channels. Nonetheless, a formal proof for an arbitrary number of transmitters $K \in \mathbb{N}$ and channels $S \in \mathbb{N}$ at a finite SNR becomes a hard task since it will require to calculate all the types of NE depending on the exact channel realizations. Hence, for the case of arbitrary parameters K, S, and SNR, we provide only numerical examples to give an insight of the general behavior. First, we evaluate the impact of the SNR for a network with a fixed number of transmitters and channels. Second, we evaluate the impact of the network load, i.e., the number of transmitters per channel for a given fixed SNR.

3.6.1 Impact of the SNR $\frac{p_{\max}}{\sigma_2}$

In Fig. 3.4 (left), we plot the network spectral efficiency as a function of the average SNR of the transmitters. Here, it is shown that in fully-loaded and over-loaded networks, i.e., $\eta = \frac{K}{S} \geqslant 1$, the gain in NSE obtained by using a discrete action set (game $\mathcal{G}_{(a)}$) increases with the SNR. Similarly, at low SNR, the NSE is the same in both cases. Conversely, for weakly loaded networks $\eta < 1$, letting all transmitters to use all the available channels is the optimal choice due to the fact that, limiting them to use only one, necessarily implies letting some channels unused.

In Fig. 3.5, we plot the probability of observing a specific number of NE in the

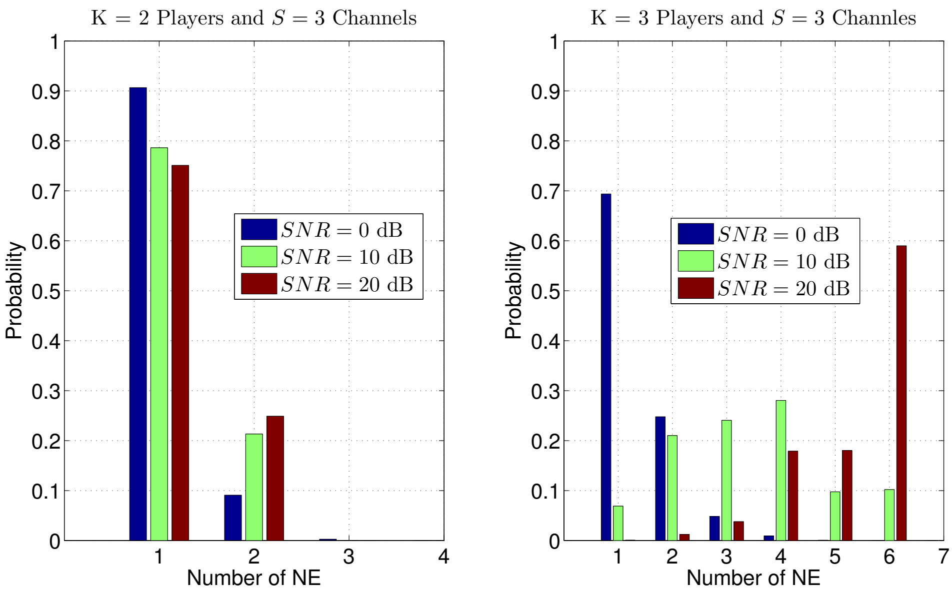

Figure 3.5: Probability of observing a specific number of NE in the game $\mathcal{G}_{(b)}$.

game $\mathcal{G}_{(b)}$ for different values of SNR. In the first case (Fig. 3.5 (left)) we consider $S = 2$ and $K = 3$, whereas in the second case (Fig. 3.5 (right)), $K = 3$ and $S = 3$. Note that from Prop. 3.4.2, the maximum number of NE is 4 and 7 in each case. However, only 3 and 6 are respectively observed with non-zero probability. Interestingly, in both cases, low SNR levels are associated with a unique NE (with high probability), whereas, high SNR levels are associated with multiple NE (with high probability).

3.6.2 Impact of Number of Transmitters per Dimension ($\frac{K}{S}$)

In Fig. 3.4 (right), we plot the NSE as a function of the number of transmitters per channel, i.e., the system load $\eta = \frac{K}{S}$. Therein, one can observe that for weakly loaded systems $\eta < 1$, playing $\mathcal{G}_{(a)}$ always leads to higher NSE than playing $\mathcal{G}_{(b)}$. On the contrary, for fully-loaded and over-loaded systems, the NSE of the game $\mathcal{G}_{(a)}$ is at least equal or better than that of the game $\mathcal{G}_{(b)}$. Interestingly, the fact that for high system loads $\eta > 2$, the NSE obtained by playing either the game $\mathcal{G}_{(a)}$ or $\mathcal{G}_{(b)}$ becomes identical implies that the analysis in Sec. 3.5.3 is also valid for the game $\mathcal{G}_{(a)}$. Finally, in Fig. 3.6, we show the fractions x_s of transmitters using channel s, with $s \in \mathcal{S}$, obtained by Monte-Carlo simulations and using (3.47) for a large network with an asymptotic ratio of players per channel equivalent to $\eta = 10$. Therein, it becomes clear that (3.47) is a precise estimation of the outcome of the non-atomic version of the game $\mathcal{G}_{(b)}$.

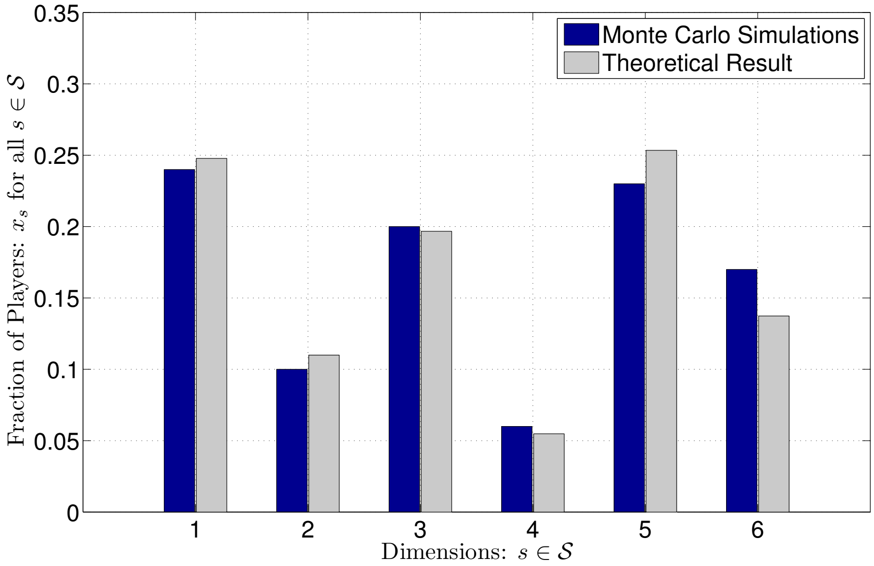

Figure 3.6: Fraction of players transmitting over channel s, with $s \in \mathcal{S}$, calculated using Monte-Carlo simulations and using Eq. (3.47) for a network with $S = 6$ channels, $\eta = 10$ players per channel, with $\mathbf{b} = (b_s)_{\forall s \in \mathcal{S}} = (0.25, 0.11, 0.20, 0.05, 0.25, 0.14)$, and $\text{SNR} = \frac{p_{\max}}{N_0 B} = 10$ dB.

3.7 Conclusions

In this chapter, it is clearly shown to what extent the equilibrium analysis of the decentralized parallel MAC differs from those conducted for other channels like Gaussian MIMO interference channels and fast fading MIMO MAC. In particular, the special structure of parallel MAC (which are important channel models in practice) and the assumption of single-user decoding at the receiver leads to the best response potential game property. The channel selection game was merely introduced in the literature but not investigated in details as it is in this chapter. In particular, a graph-theoretic interpretation is used to characterize the number of NE and the potential game property is exploited to apply learning procedures. Although all of these results are encouraging about the relevance of game-theoretic analyses of power allocation problems, important practical issues have been deliberately ignored. For example, the impact of channel estimation is not assessed at all.

Chapter 4

Learning Equilibria in Open Spectrum Sharing Games

In this chapter, we focus on the design of behavioral rules to allow radio devices to achieve an NE configuration as a result of a short interaction with its counterparts, similar to a learning process. In chapter 3, we have studied the best response dynamics (BRD) [31] and fictitious play (FP) [16]. The BRD is a behavioral rule where each radio device chooses its best transmit configuration given the configurations currently adopted by all the other devices and the current network state. Here, each radio device updates its transmit configuration either simultaneously or sequentially. In recent literature, many algorithms based on this idea have been introduced to tackle the problem of power allocation in decentralized self-configuring networks [71, 73, 80, 97, 98, 100] and have been shown to converge to NE.

In the case of FP, radio devices update their transmit configuration sequentially or simultaneously as in the BRD. Here, each radio device determines the transmit configuration which maximizes its expected performance given the empirical measures of the frequencies with which the other radio devices use each of their transmit configurations. This behavioral rule has been proved to converge to NE in several types of games relevant for wireless communications [86].

In general, the main constraint in BRD and FP and its variants is the fact that each radio device must be able to determine the (expected or instantaneous) performance obtained with each of its own transmit configurations given those adopted by all the other devices in order to choose its own optimal transmit configuration. This implies that radio devices must observe the transmit configurations of all the other devices, which is clearly a very demanding condition in practical scenarios. Nonetheless, in some network topologies and depending on the performance metric, a simple broadcast message from each receiver might be enough to implement either BRD or FP [80, 86]. However, it is not the case for most of the network topologies and performance metrics.

More advanced behavioral rules for achieving equilibria are based on reinforcement learning (RL) [18, 120]. In RL, the information required by each radio device is an observation of its own achieved performance at least every time it changes its transmit configuration. The principle of RL is as follows. After each performance observation, each radio device builds a probability distribution over all its feasible transmit configurations based on the current and all previous observations. Such probability distribution grants higher probabilities to transmit configurations leading to high performances and lower probabilities to transmit configurations leading to lower performances. At a given time, each radio device adopts a particular transmit configuration following its current probability distribution. In the wireless communication domain, this idea has been largely used and has been proved to converge to NE in some particular radio resource allocation scenarios [89, 95, 123]. The advantages of RL with respect to BRD and FP are manifold (if it converges to NE). First, the notion of synchronization in terms of round robin sequences or simultaneous transmission configuration updates is no longer relevant. This implies that each radio device can update its configuration at any time regardless of the updating timing of all the other radio devices. Second, each radio device is unaware of the presence of other radio devices and global network states. Indeed, the effect of the other radio devices' existence as well as the network state is captured by each performance observation. Note that is not the case neither in BRD nor FP, where each radio device must be able to observe the current transmit configurations of all the other devices and the current network state.

However, aside from all the attractive advantages of RL, it has an enormous drawback: each performance observation is used to directly update the probability distributions without maintaining an estimate of the performance achieved with each transmit configuration. This fact might lead the network to converge to a stationary state which is not an NE. In fact, convergence can be observed to action profiles which are suboptimal from both individual and global standpoints. We say stationary, in the sense that none of the radio devices changes its transmit configuration since it is unable to identify that other transmit configurations might bring a higher performance. In fact, any network whose normal-form game equivalent has the same structure of the classical Jordan's matching pennies game [39] and Shapley's variant of rock-scissors-paper game [103] exhibit this non-convergence effect. Consider for instance, the simple power allocation game described in [82].

In this chapter, we introduce a kind of behavioral rules which are known in the domain of Markov decision processes as actor-critic algorithms [43, 44, 107]. Here, each radio device simultaneously learns both the time-average performance achieved with each of its transmit configurations and the equilibrium probability distribution. In the domain of dynamic games, this approach was first introduced in [13] and lately in [48]. In the wireless communications domain, the idea has been recently introduced in [82]. In particular, contrary to the RL algorithms described above, whenever these behavioral rules lead to a stationary network configuration, it corresponds to a logit equilibrium (LE), which is indeed, an epsilon-close Nash equilibrium concept. In particular, we show that there exist several classes of games relevant for wireless communications where these learning dynamics always converge.

This chapter is organized as follows. In Sec. 4.1 and Sec. 4.2, the minimum requirements of a given wireless network for our approach to be valid and its corresponding game formulation are presented, respectively. In Sec. 4.3, we introduce the idea of stationary state-independent behavioral strategies, and we justify its practical relevance in the context of self-configuring networks. In Sec. 4.4, we extend the existing idea of LE to the context of stochastic games. In Sec. 4.5 and Sec. 4.6, the main contributions of the chapter is presented. First, behavioral rules allowing to learn a LE are introduced and its corresponding convergence properties are thoroughly studied. In Sec. 4.7, numerical results are presented using a multi-cell scenario where radio devices must choose the frequency band to transmit to the corresponding receiver. This chapter is concluded by Sec. 4.8.

4.1 System Model

The learning dynamics presented in this chapter apply for a large number of wireless communications scenarios. In the following, we describe the minimum conditions required to the network such that our scheme remains valid.
(i) The network is decentralized in the sense that each transmitter must autonomously determine its optimal transmission configuration. Here, a central entity able to gather complete information of the network and optimize the global network performance does not exist.
(ii) Each transmitter autonomously chooses its transmission parameters from a *finite set of choices*, e.g., power allocation polices, coding-modulation schemes or any combination of those, *etc.*, in order to optimize a given *individual performance metric*. We underline the fact that the individual performance metric of a given transmitter must depend on its own choices and those of a non-empty subset of all the other transmitters, e.g., transmission rate, bit error rate, outage probability, *etc.*
(iii) Each transmitter is able to obtain a measure of its own *instantaneous individual performance*, at least once every time it changes its transmit configuration. A practical example is the frame success rate. If the transmitter is acknowledged by the receiver frame by frame by an ACK/NACK (acknowledgment/non-acknowledgment) message, then the transmitter is able to know the instantaneous value of the number of successfully received frames.

4.2 Game Theoretic Model

The aim of this section is two-fold. First, we formulate the game which describes the wireless networks possessing the features described in the previous section. Second, we describe the desidered outcome of the corresponding game, i.e., we introduce the definition of the equilibrium we are expecting to observe in the network.

4.2.1 Game Formulation

Consider the stochastic game $\mathcal{G}$ described by the 5-tuple

$$\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{\bar{u}_k\}_{k\in\mathcal{K}}, \mathcal{H}, \{\rho_{\mathbf{h}}\}_{\mathbf{h}\in\mathcal{H}}\right). \tag{4.1}$$

The sets $\mathcal{K} = \{1, \ldots, K\}$, $\mathcal{H} = \left\{\mathbf{h}^{(1)}, \ldots, \mathbf{h}^{(H)}\right\}$ and $\mathcal{A}_k = \{A_k^{(1)}, \ldots, A_k^{(N_k)}\}$, $\forall k \in \mathcal{K}$, represent the set of players, the set of network states and the set of actions of player k, respectively. In this analysis, such sets are assumed finite, non-empty, non-unitary and time-invariant sets. In the game $\mathcal{G}$, each player represents an *active transmitter* of the network, and thus, we indifferently use the terms transmitter and player. The set of actions of a given transmitter corresponds to the set of all its feasible *transmission configurations*, for instance, power allocation policy, modulation scheme, constellation size, etc. The set $\mathcal{H}$ contains all the possible states of the network under analysis. A particular network state is assumed to be fully described by the channel realizations between all transmitters and all receivers, instantaneous quality of service requests, instantaneous energy consumption constraints, *etc.* The finiteness assumption over the set $\mathcal{H}$ might appear restrictive. However, all the network parameters are described by a finite set of information bits. Hence, in practical scenarios, there always exists a finite set of possible states.

The stochastic game $\mathcal{G}$ is played stage by stage. Each stage n lasts the time the network can be described by a given network state vector $\mathbf{h}(n) \in \mathcal{H}$, with $n \in \mathbb{N}$. In the following, without lost of generality, we assume that each state lasts a fixed time unit, e.g., the channel coherence time, and all players play at each stage. The particular network state at stage n, $\mathbf{h}(n) \in \mathcal{H}$ is a random variable. In the following of this chapter, we assume that for all stages $(n, m) \in \mathbb{N}^2$, the network states $\mathbf{h}(n)$ and $\mathbf{h}(m)$ are independent and identically distributed. Hence, it follows that $\Pr\left(\mathbf{h}(n) = \mathbf{h} \,|\mathbf{h}(m) = \mathbf{h}'\right) = \Pr\left(\mathbf{h}(n) = \mathbf{h}\right) = \rho_{\mathbf{h}}$, where $\left(\rho_{\mathbf{h}^{(1)}}, \ldots, \rho_{\mathbf{h}^{(H)}}\right) \in \triangle\left(\mathcal{H}\right)$. In particular, this assumption follows from the fact that the network state can be fully described by the channel realizations. However, in the case when the network state is determined by the channel realizations and other network parameter, e.g., a QoS request, this condition does not necessarily hold.

At each stage n, every player k adopts an action (transmission configuration), which we denote by $a_k(n) \in \mathcal{A}_k$. An action profile at stage n is a vector denoted by $\mathbf{a}(n) = (a_1(n), \ldots, a_K(n)) \in \mathcal{A}$, where

$$\mathcal{A} \stackrel{\triangle}{=} \mathcal{A}_1 \times \ldots \times \mathcal{A}_K. \tag{4.2}$$

The instantaneous performance achieved by player k is determined by the function

$$u_k : \mathcal{H} \times \mathcal{A} \to \mathbb{R}_+, \tag{4.3}$$

which measures the benefit (in the sense of Morgenstern-Neumann [68]) transmitter k obtains when it chooses a specific action $a_k(n)$ given the actions adopted by all the other transmitters $\mathbf{a}_{-k}(n)$ and the current state of the network $\mathbf{h}(n)$. At the end of each stage, player k observes a noisy sample $\tilde{u}_k(n)$ of its achieved performance $u_k\left(\mathbf{h}(n), a_k(n), \mathbf{a}_{-k}(n)\right)$. More specifically,

$$\tilde{u}_k(n) = u_k(\mathbf{h}(n), a_k(n), \mathbf{a}_{-k}(n)) + \varepsilon_{k,a_k(n)}(n), \tag{4.4}$$

where, $\forall n_k \in \{1, \ldots, N_k\}$, $\varepsilon_{k,A_k^{(n_k)}}(n)$ is the realization at time n of a random variable $\varepsilon_{k,A_k^{(n_k)}}$ which represents the additive noise on the observation of the instantaneous

performance when transmitter k plays action $A_k^{(n_k)}$. This noise can be associated to imperfect feedback or quantization distortion. Here, we assume that $\forall k \in \mathcal{K}$ and $n_k \in \{1, \ldots, N_k\}$, it holds that $\mathbb{E}\left[\varepsilon_{k,A_k^{(n_k)}}\right] = 0$.
Note that all the information gathered by player k at stage n is the 2-tuple $(a_k(n), \tilde{u}_k(n)) \in \mathcal{A}_k \times \mathbb{R}$. We denote by $\theta_k(n)$ the available information gathered by player k up to interval n, i.e.,

$$\theta_k(n) = \{(a_k(0), \tilde{u}_k(0)), \ldots, (a_k(n-1), \tilde{u}_k(n-1))\}. \tag{4.5}$$

We refer to $\theta_k(n)$ as the private history of player k at time n. The set of all possible private histories of player k at time n is denoted by $\Theta_k(n)$, and,

$$\Theta_k(n) = (\mathcal{A}_k \times \mathbb{R})^n . \tag{4.6}$$

The set of all possible private histories of player k in the infinite play is denoted by

$$\Theta_k = (\mathcal{A}_k \times \mathbb{R})^{\mathbb{N}} . \tag{4.7}$$

Now, using the above definition of private history, we introduce the idea of behavioral strategy (BS). A BS of player k is an infinite sequence of functions

$$\sigma_k = \{\sigma_{k,n}\}_{n>0} , \tag{4.8}$$

such that for all $n > 0$, the function

$$\sigma_{k,n} : \Theta_k(n) \to \triangle(\mathcal{A}_k) \tag{4.9}$$

determines the probability distribution with which player k takes the action $a_k(n)$ given the private history $\theta_k(n)$. Denote by Σ_k, the set of all possible BS of player k and let $\Sigma = \Sigma_1 \times \ldots \times \Sigma_K$ be the set of all BS profiles.
Following this idea, given any behavioral strategy $\sigma = (\sigma_1, \ldots, \sigma_K)$, the initial action profile $\mathbf{a}(0) \in \mathcal{A}$ and the corresponding observations $\tilde{u}_1(0), \ldots, \tilde{u}_K(0)$ induce a set of sequences of probability distributions $\{\pi_k(n)\}_{n>0}$, for all $k \in \mathcal{K}$, where

$$\pi_k(n) = \left(\pi_{k,A_k^{(1)}}(n), \ldots, \pi_{k,A_k^{(N_k)}}(n)\right) \in \triangle(\mathcal{A}_k) \tag{4.10}$$

and $\forall n_k \in \{1, \ldots, N_k\}$, $\pi_{k,A_k^{(n_k)}}(n)$ represents the probability that player k plays action $A_k^{(n_k)} \in \mathcal{A}_k$ at time n, i.e.,

$$\pi_{k,A_k^{(n_k)}}(n) = \Pr\left(a_k(n) = A_k^{(n_k)}\right). \tag{4.11}$$

Hence the set of sequences $\{\pi_k(n)\}_{n>0}$ induced by $\mathbf{a}(0) \in \mathcal{A}$ and the vector $\tilde{\mathbf{u}}(0) = (\tilde{u}_1(0), \ldots, \tilde{u}_K(0))$, together with the initial probability distributions $\pi(0) = (\pi_1(0), \ldots, \pi_K(0))$ induce a probability distribution over all the possible sequences of action profiles $\{\mathbf{a}(0), \mathbf{a}(1), \ldots, \}$. We denote the expectation with respect to such probability distribution by $\mathbb{E}_{(\mathbf{a}(0),\tilde{\mathbf{u}}(0),\sigma)}$. Then, given the available information for player k, its long-term expected performance can be measured by the function, $\bar{u}_k : \Sigma_1 \times \ldots \times \Sigma_K \to \mathbb{R}_+$, where,

$$\bar{u}_k(\sigma_k, \sigma_{-k}) = \lim_{n\to\infty} \mathbb{E}_{(\mathbf{a}(0),\tilde{\mathbf{u}}(0),\sigma)}\left[\frac{1}{n}\sum_{i=1}^{n} \tilde{u}_k(i)\right]. \tag{4.12}$$

The function (4.12) captures the situation in which the interaction between all transmitters in the network lasts many time intervals and the instantaneous performance is insignificant as compared with the performance in all the other time intervals.
In the following, the game $\mathcal{G}$ is analyzed assuming that the aim of each player k is to choose a BS $\sigma_k \in \Sigma_k$ such that it maximizes its performance metric (4.12) given the BS $\sigma_{-k} \in \Sigma_{-k}$ adopted by all the other players. In particular, we look for a BS profile $\sigma^* = (\sigma_1^*, \ldots, \sigma_K^*) \in \Sigma_1 \times \ldots \Sigma_K$ such that none of the players can obtain a performance improvement by unilaterally using other BS. We provide a more precise concept of this expected solution of the stochastic game $\mathcal{G}$ in the following subsection.

4.2.2 Nash Equilibrium and ϵ-Equilibrium

In the following, we describe the concept of ϵ-equilibrium and Nash equilibrium in the context of the stochastic game $\mathcal{G}$ in (4.1). First, let us define the ϵ-equilibrium as follows,

Definition 4.2.1 (ϵ-Equilibrium in the game $\mathcal{G}$) *Let $\epsilon > 0$. In the game $\mathcal{G}$, a strategy profile $\boldsymbol{\sigma}^* \in \Sigma_1 \times \ldots \times \Sigma_K$ is an ϵ-equilibrium if it satisfies, for all $k \in \mathcal{K}$ and for all $\sigma_k \in \Sigma_k$, that*

$$\bar{u}_k(\sigma_k^*, \sigma_{-k}^*) \geqslant \bar{u}_k(\sigma_k, \sigma_{-k}^*) - \epsilon, \tag{4.13}$$

independently of the initial action profile and observation vector, $\mathbf{a}(0) \in \mathcal{A}$ and $\tilde{\mathbf{u}}(0) \in \mathbb{R}^K$, respectively.

An ϵ-equilibrium can be interpreted as a BS profile such that, none of the players can obtain an improvement superior to ϵ by unilaterally changing its own BS. Note also that by letting $\epsilon = 0$ in Def. 4.2.1, the classical definition of Nash equilibrium is obtained.
In the following section, we discuss the feasibility of achieving these equilibrium concepts in the game $\mathcal{G}$.

4.3 A Note on Stationary State-Independent Behavioral Strategies

The aim of this section is twofold. First, it aims to introduce the concept of stationary state independent behavioral strategies (SSI-BS), which are considered the simplest class of behavioral strategies (BS) in stochastic games [70, 105]. Second, it aims to identify the expected performance which can be achieved using both BS and SSI-BS. We conclude this section by presenting the justification for restricting the analysis of the game $\mathcal{G}$ to the set of SSI-BS.

4.3.1 Stationary State Independent Behavioral Strategies (SSI-BS)

Let the set of stationary state independent behavioral strategy (SSI-BS) profiles be denoted by $\bar{\Sigma}$ and let a given SSI-BS profile be defined as follows,

Definition 4.3.1 (Stationary State Independent BS) *Consider the game $\mathcal{G}$ and let $\sigma \in \Sigma$ be a behavioral strategy. Then, σ is said to be stationary state-independent (SSI) if for all $k \in \mathcal{K}$ and any two private histories $\theta_k(n) \in \mathbf{\Theta}_k(n)$ and $\theta_k(m) \in \mathbf{\Theta}_k(m)$, with $n \neq m$, it follows that*

$$\sigma_{k,n}\left(\theta_k(n)\right) = \sigma_{k,m}\left(\theta_k(m)\right), \tag{4.14}$$

independently of the states $\mathbf{h}(n)$ and $\mathbf{h}(m)$.

From Def. 4.3.1, it can be implied that for a player k, a SSI-BS does not depend on any of the previous actions $a_k(0), \ldots, a_k(n-1)$ and neither on the previous nor current states $\mathbf{h}(0), \ldots, \mathbf{h}(n)$. Thus, a SSI-BS $\sigma_k \in \bar{\Sigma}$ can be identified by a vector $\pi = (\pi_1, \ldots, \pi_K) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$, such that, $\forall k \in \mathcal{K}$ and $\forall\left(\theta_k(n), \mathbf{h}(n)\right) \in \Theta_k(n) \times \mathcal{H}$, it holds that $\sigma_{k,n}(\theta_k(n)) = \pi_k$. In the following, we indifferently use the infinite set of sequences $\sigma_k = \{\pi_k(n) = \pi_k\}_{n>0}$ or the vectors π_k, with $k \in \mathcal{K}$, to refer to the SSI-BS σ. Moreover, with a slight abuse of notation, we indifferently write either $\bar{u}_k\left(\sigma_k, \sigma_{-k}\right)$ or $\bar{u}_k\left(\pi_k, \pi_{-k}\right)$ to denote the achieved performance of player k given the SSI-BS σ.

4.3.2 Achievable Performance with SSI-BS

A vector of individual performance $\mathbf{r} = (r_1, \ldots, r_K) \in \mathbb{R}^K$ is achievable in the game $\mathcal{G}$ with the set of behavioral strategies Σ if there exists a strategy profile $\sigma \in \Sigma$ and an initialization $(\mathbf{a}(0), \hat{\mathbf{u}}(0)) \in \mathcal{A} \times \mathbb{R}_+^K$, such that for all $k \in \mathcal{K}$, $r_k = \bar{u}_k\left(\sigma_k, \sigma_{-k}\right)$. In order to determine the set of achievable individual performance vectors with the set of BS Σ and the set of SSI-BS $\bar{\Sigma}$, we first introduce the concept of asymptotic average strategic behavior (AASB) [120].

Definition 4.3.2 (Asymptotic Average Strategic Behavior (AASB)) *Consider the game $\mathcal{G}$ and let $(\mathbf{a}(0), \hat{\mathbf{u}}(0)) \in \mathcal{A} \times \mathbb{R}_+^K$ and $\sigma \in \Sigma$ be an initialization and a strategy profile, respectively. Assume that $(\mathbf{a}(0), \hat{\mathbf{u}}(0))$ together with σ induce the infinite sequence of probability distributions $\{\pi_k(n)\}_{n>0}$, for all $k \in \mathcal{K}$, such that for all $n \in \mathbb{N}$, $\pi_k(n) \in \triangle(\mathcal{A}_k)$. Then, the asymptotic average strategic behavior associated with the triplet $(\mathbf{a}(0), \hat{\mathbf{u}}(0), \sigma)$ is a probability distribution $\kappa^* = (\kappa^*_{\mathbf{a}})_{\forall \mathbf{a} \in \mathcal{A}} \in \triangle(\mathcal{A})$, such that, $\forall \mathbf{a} \in \mathcal{A}$,*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \prod_{j=1}^{K} \pi_{j,a_j}(n) = \kappa^*_{\mathbf{a}}. \tag{4.15}$$

The Def. 4.3.2 leads us to the following lemma.

Lemma 4.3.3 *Consider the game $\mathcal{G}$ and let $(\mathbf{a}(0), \hat{\mathbf{u}}(0)) \in \mathcal{A} \times \mathbb{R}_+$ and $\sigma^* \in \Sigma$ be an initialization and a BS profile, respectively. Assume that the asymptotic average strategic behavior induced by $(\mathbf{a}(0), \hat{\mathbf{u}}(0), \sigma^*)$ is $\kappa^* \in \triangle(\mathcal{A})$. Then, the performance achieved by player k, with $k \in \mathcal{K}$, is*

$$\bar{u}_k\left(\sigma_k^*, \sigma_{-k}^*\right) = \sum_{\mathbf{a}\in\mathcal{A}} \sum_{\mathbf{h}\in\mathcal{H}} u_k(\mathbf{h}, a_k, \mathbf{a}_{-k}) \rho_{\mathbf{h}} \kappa_{\mathbf{a}}^*. \tag{4.16}$$

The proof of Lemma 4.3.3 is presented in App. H. The relevance of Lemma 4.3.3 stems from the fact that it allows to express the performance achievable with the behavioral strategy σ^* and the initialization pair $(\mathbf{a}(0), \hat{\mathbf{u}}(0))$ by a static probability distribution which coincides with the corresponding AASB κ^* (Def. 4.3.2). This implies that it might exist a loss of optimality by studying the equilibria of the game $\mathcal{G}$ only on the set of SSI-BS. More precisely, the following can be stated.

Proposition 4.3.1 *Consider the game $\mathcal{G}$ and let $(\mathbf{a}(0), \hat{\mathbf{u}}(0)) \in \mathcal{A} \times \mathbb{R}_+$ and $\sigma^* \in \Sigma$ be an initialization and a strategy profile, respectively. Denote by $\kappa^* \in \triangle(\mathcal{A})$ the AASB induced by the triplet $(\mathbf{a}(0), \hat{\mathbf{u}}(0), \sigma^*)$. Then, there exists a static state-independent behavioral strategy $\sigma' \in \bar{\Sigma}$ which satisfies*

$$\bar{u}_k(\sigma_k^*, \sigma_{-k}^*) = \bar{u}_k(\sigma_k', \sigma_{-k}'), \tag{4.17}$$

if and only if there exists a vector $\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$, such that, for all $\mathbf{a} \in \mathcal{A}$, $\kappa_{\mathbf{a}}^* = \prod_{k=1}^{K} \pi_{k,a_k}^*$.

The proof of Prop. 4.3.1 is immediately from Lemma 4.3.3. The relevance of this proposition stems from the fact that it proves that any expected performance vector $\mathbf{r}$ of the game $\mathcal{G}$ is achievable with both BS and SSI-BS, as long as the corresponding AASB is a product distribution, i.e., $\forall \mathbf{a} \in \mathcal{A}$, $\kappa_{\mathbf{a}}^* = \prod_{k=1}^{K} \pi_{k,a_k}$. Behavioral strategies leading to a correlated AASB, e.g., regret matching, are described in [120], [35]. In the following, we focus the analysis of the game $\mathcal{G}$ on the set of SSI-BS.

4.3.3 Pertinence of the SSI-BS

The motivations for restricting the set of BS to the set of SSI-BS are several. First, note that none of the players is able to identify the current network state at each stage of the game. Moreover, given the information gathered by player k up to stage $n-1$, i.e., $\theta_k(n-1)$, with $k \in \mathcal{K}$ and $n > 1$, it is not possible to infer any information about the network state $\mathbf{h}(n)$. This implies that, player k is unable to calculate an optimal probability distribution $\pi_k(n)$ at each stage n, since the ignorance of the network state at each stage implies the ignorance of a closed form expression of the instantaneous performance metric u_k and the long-term performance metric $\bar{u}_k$. Thus, regardless of the stage n and all the information gathered up to such stage n, all players face the same scenario.

A second justification can be given considering that at a given stage, a behavioral strategy might condition the determination of the probability distribution $\pi_k(n) \in \triangle(\mathcal{A}_k)$ on the entire history $\theta_k(n)$, which from an implementation point of view is a huge task. Moreover, it might require a massive amount of memory for storing $\theta_k(n)$ as long as $n \to \infty$.
In the following of this chapter, we restrict the analysis of equilibria of the game $\mathcal{G}$ to the set of SSI-BS and we accept the eventual loss of performance which implies the non-correlation between the individual actions.

4.4 Logit Equilibrium

The concept of logit equilibrium (LE) is a particular case of a more general class of equilibria known as quantal response equilibria introduced by McKelvey and Palfrey in the context of strategic games in normal form [57] and extensive form [58]. The concept of logit equilibrium and the ideas on which it relies on have many interpretations. The interested reader is referred to [49, 56–58, 120] for more general discussions.
In this section, we adapt the concept of LE to the game $\mathcal{G}$. Here, we do not give any particular interpretation to the LE other than the simple idea of ϵ-equilibrium described in Def. 4.2.1. In particular, we discuss the existence and uniqueness of the LE in the game $\mathcal{G}$ in the set of SSI-BS.

4.4.1 Logit Equilibrium in SSI-BS

Before we provide a formal definition of the logit equilibrium, we introduce the idea of logit best response. For doing so, consider the game $\mathcal{G}$ and assume that player k is able to obtain an estimate of its own performance (4.3) in the hypothetical case it played the same action $A_k^{(n_k)}$, with $n_k \in \{1, \ldots, N_k\}$, during the whole realization of the game, while all the other players play the SSI-BS $\pi_j \in \triangle(\mathcal{A}_j)$, with $j \in \mathcal{K} \setminus \{k\}$, respectively. More specifically, assume that each player $k \in \mathcal{K}$ is able to calculate the vector,

$$\bar{\mathbf{u}}_k(\cdot, \pi_{-k}) = \left(\bar{u}_k\left(\mathbf{e}_k^{(1)}, \pi_{-k}\right), \ldots, \bar{u}_k\left(\mathbf{e}_k^{(N_k)}, \pi_{-k}\right)\right), \tag{4.18}$$

for all $\pi_{-k} \in \triangle(\mathcal{A}_{-k})$. Then, the logit best response can be defined as follows.

Definition 4.4.1 (Logit Best Response) *Consider the game $\mathcal{G}$ and let the vector $\pi_{-k} \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_{k-1}) \times \triangle(\mathcal{A}_{k+1}) \times \ldots \times \triangle(\mathcal{A}_K)$ represent a given SSI-BS profile, with $k \in \mathcal{K}$. Then, the logit best response of player k, with parameter $\gamma_k > 0$, is the probability distribution $\beta_k^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})) \in \triangle(\mathcal{A}_k)$ such that, $\beta_k^{(\gamma_k)} : \mathbb{R}^{N_k} \to \triangle(\mathcal{A}_k)$ is the logit function,*

$$\beta_k^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})) = \left(\beta_{k,A_k^{(1)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k})), \ldots, \beta_{k,A_k^{(N_k)}}^{(\gamma_k)}(\bar{\mathbf{u}}_k(\cdot, \pi_{-k}))\right) \tag{4.19}$$

and $\forall n_k \in \{1, \ldots, N_k\}$,

$$\beta^{(\gamma_k)}_{k,A_k^{(n_k)}}\left(\bar{\mathbf{u}}_k(\cdot,\pi_{-k})\right) = \frac{\exp\left(\gamma_k \bar{u}_k(\mathbf{e}_k^{(n_k)},\pi_{-k})\right)}{\sum_{m=1}^{N_k}\exp\left(\gamma_k \bar{u}_k(\mathbf{e}_k^{(m)},\pi_{-k})\right)}. \tag{4.20}$$

Note Def. 4.4.1 does not state anything about the selection of the parameter γ_k for player k. The impact of this parameter is studied later in this chapter. From Def. 4.4.1, it can be implied that at each stage of the game, every action of a given player has a (stationary state-independent) non-zero probability of being played, i.e., $\forall k \in \mathcal{K}$ and $\forall n_k \in \{1, \ldots, N_k\}$ and $\forall \gamma_k \in \mathbb{R}_+$, it holds that, $\beta^{(\gamma_k)}_{k,A_k^{(n_k)}}\left(\bar{\mathbf{u}}_k\left(\cdot,\pi_{-k}\right)\right) > 0$. More generally, it can be stated that the logit best response in SSI-BS is represented by a probability distribution that assigns high probabilities to the actions associated to a high average performance and low probability to actions associated to low average performance. For instance, let exist a couple $(m,n) \in \{1,\ldots,N_k\}^2$, such that $\bar{u}_k(\mathbf{e}_k^{(n)},\pi_{-k}) < \bar{u}_k(\mathbf{e}_k^{(m)},\pi_{-k})$, then $0 < \beta^{(\gamma_k)}_{k,A_k^{(n)}}\left(\bar{\mathbf{u}}_k\left(\cdot,\pi_{-k}\right)\right) < \beta^{(\gamma_k)}_{k,A_k^{(m)}}\left(\bar{\mathbf{u}}_k\left(\cdot,\pi_{-k}\right)\right)$. Finally, note that conversely to the case of the best response in the case of Nash equilibrium [66], the logit best response of player k is unique for all the SSI-BS profiles the other players might adopt.
Using Def. 4.4.1, we define the logit equilibrium as follows,

Definition 4.4.2 (Logit Equilibrium in SSI-BS) *Consider the game* $\mathcal{G}$ *and let the vector* $\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ *represent a stationary state-independent behavioral strategy (SSI-BS). Then,* π^* *is a logit equilibrium SSI-BS profile with parameter* $\gamma = (\gamma_1, \ldots, \gamma_K)$ *if for all* $k \in \mathcal{K}$*, it holds that,*

$$\pi_k^* \;=\; \beta_k^{(\gamma_k)}\left(\bar{u}_k\left(\mathbf{e}_1^{(N_k)},\pi_{-k}^*\right),\ldots,\bar{u}_k\left(\mathbf{e}_{N_k}^{(N_k)},\pi_{-k}^*\right)\right). \tag{4.21}$$

At a logit equilibrium, each player plays during the whole game realization, the logit best response (Def. 4.4.1) to all the other players SSI-BS. This implies that, at some given game stages, the actions taken by player k do not maximize the instantaneous performance u_k and thus, other strategy profile might bring a higher long-term performance $\bar{u}_k$. In the following proposition, we determine the impact of playing actions which are not performance-maximizers.

Proposition 4.4.1 *Let the vector* $\pi^* \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ *be a LE in SSI-BS of the game* $\mathcal{G}$*, with parameter* $\gamma = (\gamma_1, \ldots, \gamma_K)$*. Hence,* $\forall k \in \mathcal{K}$ *and* $\forall \pi_k' \in \triangle(\mathcal{A}_k)$*, it holds that,*

$$\bar{u}_k\left(\pi_k^*,\pi_{-k}^*\right) \geqslant \bar{u}_k\left(\pi_k',\pi_{-k}^*\right) - \frac{1}{\gamma_k}\ln\left(\frac{1}{N_k}\right). \tag{4.22}$$

Thus, π^* *is an* ϵ*-equilibrium (Def. 4.2.1) and* $\epsilon = \max_{k\in\mathcal{K}}\left(\frac{1}{\gamma_k}\ln\left(N_k\right)\right)$.

The Prop. 4.4.1 is a well known result [120] and has been included here for the sake of completeness. From Prop. 4.4.1, it becomes clear that the maximum loss of performance player k might experience is bounded by $\frac{1}{\gamma_k}\ln(N_k)$. In the following subsections we discuss the existence and uniqueness of the LE in the game $\mathcal{G}$.

4.4.2 Existence in SSI-BS

It is important to remark that Def. 4.4.2 implies a fixed point equation. For instance, let $\zeta : \triangle(\mathcal{A}_1) \times, \ldots \times \triangle(\mathcal{A}_K) \rightarrow \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ be defined as follows,

$$\zeta(\pi) = \left(\beta_1^{(\gamma_1)}\left(\bar{\mathbf{u}}_1\left(\cdot, \pi_{-1}\right)\right), \ldots, \beta_K^{(\gamma_K)}\left(\bar{\mathbf{u}}_K\left(\cdot, \pi_{-K}\right)\right)\right). \tag{4.23}$$

Then, if π^* is a logit equilibrium it holds that $\pi^* = \zeta(\pi^*)$. This observation leads to the following result (Theorem 1 in [57]),

Theorem 4.4.3 (Existence of the LE) *The stochastic game $\mathcal{G}$ has at least one logit equilibrium in the set of stationary state-independent behavioral strategies.*

The proof of Theorem 4.4.3 relies on the fact that the multi-dimensional function ζ is continuous in $\triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$. Thus, by Brouwer's fixed point theorem, there exists at least one π^* such that $\pi^* = \zeta(\pi^*)$.

4.4.3 Uniqueness in SSI-BS

The uniqueness of the LE in the game $\mathcal{G}$ in SSI-BS is strongly related to the parameters γ_k, with $k \in \mathcal{K}$. For instance, when $\forall k \in \mathcal{K}$, $\gamma_k \rightarrow 0$, there exits a unique LE in SSI-BS and corresponds to the vectors $\pi_k = \frac{1}{N_k}(1, \ldots, 1) \in \triangle(\mathcal{A}_k)$. This LE is unique, independently of the number of NE the game $\mathcal{G}$ might possess. Conversely, when $\forall k \in \mathcal{K}$, $\gamma_k \rightarrow \infty$, the set of LE becomes identical to the set of NE in pure strategies and thus, the game $\mathcal{G}$ exhibits as many LE as NE in pure strategies might exist in $\mathcal{G}$. More precisely, let the set $\triangle_{\mathrm{LE}}^{(\gamma)} \subset \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ contain all the vectors representing SSI-BS profiles leading to an LE in the game $\mathcal{G}$, i.e.,

$$\triangle_{\mathrm{LE}}^{(\gamma)} = \left\{\pi \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K) : \forall k \in \mathcal{K},\ \pi_k = \beta_k^{(\gamma_k)}\left(\bar{\mathbf{u}}_k\left(\cdot, \pi_{-k}\right)\right)\right\}. \tag{4.24}$$

Hence, from Theorem 4.4.3, it holds that $\forall \gamma \in \mathbb{R}_+^K$, $\left|\triangle_{\mathrm{LE}}^{(\gamma)}\right| > 0$. Other properties of the set $\triangle_{\mathrm{LE}}^{(\gamma)}$ are described by the following theorem (Theorem 3 in [58]).

Theorem 4.4.4 (Properties of the LE) *Consider the stochastic game $\mathcal{G}$. Then,*

- $\left|\triangle_{\mathrm{LE}}^{(\gamma)}\right|$ *is odd number for almost all $\gamma \in \mathbb{R}_+^K$.*
- *The graph of $\triangle_{\mathrm{LE}}^{(\gamma)}$ contains a unique branch which starts at the centroid, for $\gamma = (0, \ldots, 0) \in \mathbb{R}^K$, and converges to a unique NE as $\forall k \in \mathcal{K}$, $\gamma_k \rightarrow \infty$.*

The first property in Theorem 4.4.4 states that the number of LE in the game $\mathcal{G}$ is odd. Interestingly, the same property has been claimed for the NE [117]. Property two is in particular very interesting since it implies that the concept of LE can be used as an NE selection method if the parameters γ_k, for all $k \in \mathcal{K}$, are left to be time-variant. However, we do not exploit this property and we assume the parameters γ_k are fixed.
In the next section, we study behavioral strategies which allows transmitters to achieve an LE for a given set of fixed parameters $\gamma_1, \ldots, \gamma_K$.

4.5 Learning Logit Equilibria

In this section, we design behavioral strategy profiles $\sigma = (\sigma_1, \ldots, \sigma_K) \in \Sigma$ such that given the information gathered by player k at each stage n, i.e., given the sets $\{\theta_k(n)\}_{n>0}$ for all $k \in \mathcal{K}$, it is able to generate infinite sequences $\{\pi_k(n)\}_{n>0}$, such that, $\lim_{n\to\infty} ||\pi_k(n) - \pi_k^*|| = 0$, where $\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ is a logit equilibrium in SSI-BS of the game $\mathcal{G}$ (Def. 4.4.2). Our interest in this kind of strategies is justified by the following proposition.

Proposition 4.5.1 *Consider the game $\mathcal{G}$ and let the vector $\pi^* = (\pi_1^*, \ldots, \pi_K^*) \in \triangle(\mathcal{A}_1)\times\ldots\times\triangle(\mathcal{A}_K)$ be a SSI-BS profile corresponding to a logit equilibrium. Then, any BS profile $\sigma \in \Sigma$ which generates infinite sequences $\{\pi_k(n)\}_{n>0}$, with $k \in \mathcal{K}$, for all $\mathbf{a}(0) \in \mathcal{A}$ and*

$$\lim_{n\to\infty} ||\pi_k(n) - \pi_k^*|| = 0, \tag{4.25}$$

is an ϵ-equilibrium of the game $\mathcal{G}$.

The proof of Prop. 4.5.1 follows from Def. 4.2.1. From Prop. 4.5.1, it can be implied that any behavioral strategy satisfying (4.25) is an ϵ-equilibrium and achieves the same performance corresponding to one of the logit equilibrium in SSI-BS discussed in the previous section.

The remaining of this section is divided in three parts. In the first part, we study learning processes aiming to calculate the expected performance obtained by player k with each of its actions given its available information $\{\theta_k(n)\}_{n>0}$. That is, the vector $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ (described in (4.18)) at each game stage n. The relevance of letting player k to learn the vector $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ stems from the fact that it allows it calculate the logit best response at each time n. Second, we study the learning processes aiming to learn the probability distribution corresponding to a logit equilibrium. In this part, the underlying assumption is that player k possesses perfect knowledge of the vector $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$, with $k \in \mathcal{K}$. Finally, we join the previous results and we introduce a family of behavioral strategies which are ϵ-equilibrium in the game $\mathcal{G}$. Such families of strategies achieve the performance corresponding to at least one of the logit equilibria of the game $\mathcal{G}$ in SSI-BS for a given constant parameter $\gamma = (\gamma_1, \ldots, \gamma_K) \in \mathbb{R}^K$.

4.5.1 Learning the Expected Performance

Let $T_{k,A_k^{(n_k)}}(n) \in \mathbb{N}$, with $k \in \mathcal{K}$ and $n_k \in \{1, \ldots, N_k\}$, be the number of times that player k has played action $A_k^{(n_k)}$ up to the game stage $n \in \mathbb{N}$, i.e.,

$$T_{k,A_k^{(n_k)}}(n) = \sum_{s=0}^{n-1} \mathbb{1}_{\left\{a_k(s)=A_k^{(n_k)}\right\}}. \tag{4.26}$$

Let also $\hat{\pi}_{k,A_k^{(n_k)}}(n)$ be the empirical frequency with which player k has used action $A_k^{(n_k)}$ up to time n, i.e.,

$$\hat{\pi}_{k,A_k^{(n_k)}}(n) = \frac{T_{k,A_k^{(n_k)}}(n)}{n}. \tag{4.27}$$

Now, consider the following assumption.
(A0) *For all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, N_k\}$, it holds that*

$$\lim_{n\to\infty} \hat{\pi}_{k,A_k^{(n_k)}}(n) = \pi^+_{k,A_k^{(n_k)}}, \tag{4.28}$$

where $\pi_k^+ = \left(\pi^+_{k,A_k^{(1)}}, \ldots, \pi^+_{k,A_k^{(N_k)}}\right) \in \triangle(\mathcal{A}_k)$ *and* $\pi^+_{k,A_k^{(n_k)}} > 0$.

Let the N_k-dimensional vector $\hat{\mathbf{u}}_k(n) = \left(\hat{u}_{k,A_k^{(1)}}(n), \ldots, \hat{u}_{k,A_k^{(N_k)}}(n)\right)$ be the estimation at the game stage n that player k possesses of the vector $\bar{\mathbf{u}}_k(\cdot, \pi^+_{-k}) = \left(\bar{u}_k\left(\mathbf{e}_k^{(1)}, \pi^+_{-k}\right), \ldots, \bar{u}_k\left(\mathbf{e}_k^{(N_k)}, \pi^+_{-k}\right)\right)$ described in (4.18). Hence, assuming that assumption $(A0)$ holds, the main result of this subsection is presented in the following lemma.

Lemma 4.5.1 (Learning the Achieved Utility) *Consider the game $\mathcal{G}$ and let assumption $(A0)$ hold. Then, $\forall k \in \mathcal{K}$ and $\forall n_k \in \{1, \ldots, N_k\}$ and $n \in \mathbb{N}$, the time-averaging of the observations $\tilde{u}_k(n)$, i.e.,*

$$\hat{u}_{k,A_k^{(n_k)}}(n) = \frac{1}{T_{k,A_k^{(n_k)}}(n)} \sum_{s=0}^{n-1} \tilde{u}_k(s) \mathbb{1}_{\left\{a_k(s)=A_k^{(n_k)}\right\}}, \tag{4.29}$$

or any iterative algorithm of the form,

$$\begin{aligned} \hat{u}_{k,A_k^{(n_k)}}(n) &= \hat{u}_{k,A_k^{(n_k)}}(n-1) + \\ & \alpha_k(n) \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{\pi^+_{k,A_k^{(n_k)}}} \left(\tilde{u}_k(n-1) - \hat{u}_{k,A_k^{(n_k)}}(n-1)\right), \end{aligned} \tag{4.30}$$

where,

$$\lim_{T\to\infty} \sum_{n=0}^{T} \alpha_k(n) = +\infty \quad \text{and} \quad \lim_{T\to\infty} \sum_{n=0}^{T} \alpha_k(n)^2 < +\infty, \tag{4.31}$$

satisfies that,

$$\lim_{n\to\infty} \hat{u}_{k,A_k^{(n_k)}}(n) = \bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi^+_{-k}). \tag{4.32}$$

The proof of Lemma 4.5.1 is presented in Appendix I. It is important to remark that under assumption $(A0)$, both the time-average (4.29) or the iterative algorithm (4.30) converge asymptotically to the expected performance in (4.32). However, there exists a difference in the convergence time.

Theorem 4.5.2 *Consider the game $\mathcal{G}$ and let the assumption $(A0)$ hold. For all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, N_k\}$, denote by $\eta_{k,A_k^{(n_k)}}(n)$ the performance estimation error of transmitter k with respect to the action $A_k^{(n_k)}$ at game stage $n \in \mathbb{N}$, i.e.,*

$$\eta_{k,A_k^{(n_k)}}(n) = \left|\hat{u}_{k,A_k^{(n_k)}}(n) - \bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi^*_{-k}\right)\right|. \tag{4.33}$$

Let $\eta > 0$ be an error threshold and denote by $m^{\dagger}_{k,A_k^{(n_k)}} \in \mathbb{N}$ and $m^{+}_{k,A_k^{(n_k)}} \in \mathbb{N}$ two stage indices such that $\forall m > m^{\dagger}_{k,A_k^{(n_k)}}$, $\eta_{k,A_k^{(n_k)}}(m) < \eta$, when $\hat{u}_{k,A_k^{(n_k)}}$ follows the rule (4.29) and $\forall m > m^{+}_{k,A_k^{(n_k)}}$, $\eta_{k,A_k^{(n_k)}}(m) < \eta$ when $\hat{u}_{k,A_k^{(n_k)}}$ follows the rule (4.30). Then, the following holds for a sufficiently small η,

$$m^{\dagger}_{k,A_k^{(n_k)}} \geqslant \frac{1}{\pi^{+}_{k,A_k^{(n_k)}}} \ln\left(\frac{\left|\bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^{+}_{-k}\right) - \hat{u}_{k,A_k^{(n_k)}}(0)\right|}{\eta}\right), \tag{4.34}$$

$$m^{+}_{k,A_k^{(n_k)}} \geqslant \ln\left(\frac{\left|\bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^{+}_{-k}\right) - \hat{u}_{k,A_k^{(n_k)}}(0)\right|}{\eta}\right). \tag{4.35}$$

The proof of Theorem 4.5.2 is presented in the appendix J. From Theorem 4.5.2 it can be easily concluded that the learning dynamics given by (4.30) converges faster than the time-averaging rule (4.29) under assumption $(A0)$.

4.5.2 Learning the Equilibrium Probability Distributions

In this subsection, we focus on the design of the sequences $\{\pi_k(n)\}_{n>0}$, with $k \in \mathcal{K}$, which correspond to a ϵ-equilibrium according to Prop. 4.5.1. As a first step in this direction, we consider the following assumption.

(B0) *For all $k \in \mathcal{K}$, at each time interval $n \in \mathbb{N}$, a perfect estimation of the expected performance of player k can be obtained, i.e., $\forall n_k \in \{1, \ldots, N_k\}$, $\hat{u}_{k,A_k^{(n_k)}}(n) = \bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi_{-k}(n)\right)$.*

An intuitive interpretation of the n_k-th component of the vector $\hat{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$, at a given game stage n, is the following. The value $\hat{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi_{-k}(n)\right)$ represents the expected performance of player k when it has played the action $A_k^{(n_k)}$ during the whole game realization while the other players have played a SSI-BS represented by the vector $\pi_{-k}(n)$. Thus, given the vector $\hat{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$, the optimal probability distribution at time n is the one which assigns unit probability to the action associated to the maximum expected utility. This behavioral strategy is known as fictitious play [16] and it is proved to converge to Nash equilibria in several classes of games [120]. However, it is not compatible to the assumption $(A0)$, since it assigns zero probability to the actions associated to the lowest expected utility. This implies that actions associated with low expected performance are never played, and thus, its associated expected performance cannot be estimated. An ϵ-optimal behavioral strategy is the one which generates the probability distributions $\{\beta_k^{(\gamma_k)}\left(\hat{\mathbf{u}}_k(n)\right)\}_{n>0}$ in (4.19). This behavioral strategy is known as smoothed fictitious play [41], [9] and it is in line with assumption $(A0)$. This behavioral rule is proved to converge to ϵ-equilibrium in several classes of games [9] under assumption $(B0)$. In the following, we study the behavioral strategy which generates the sequence of probability distributions $\{\pi_k(n)\}_{n>0}$, for all $k \in \mathcal{K}$, where the probability distribution of player

k at time n is given by the following stochastic approximation algorithm (SAA),

$$\pi_k(n) = \pi_k(n-1) + \lambda_k(n)\left(\beta_k^{(\gamma_k)}\left(\hat{\mathbf{u}}_k(n)\right) - \pi_k(n-1)\right). \tag{4.36}$$

Here, λ_k is a learning rate which is described later on in this section. As we shall see, the motivation of using the SAA in (4.36) is the fact that it allows to control the variation that player k makes of its probability distribution $\pi_k(n)$ from the game stage n to game stage $n+1$ by modifying the value of λ_k. Note that as $\lambda_k \to 0$, player k tends to be conservative. Conversely, when $\lambda_k \to 1$, player k approaches the smoothed fictitious play.

The main result of this subsection considering the assumption $(B0)$ is presented in the following lemma.

Lemma 4.5.3 (Learning the Strategy Profile) *Consider the game $\mathcal{G}$ and assume that assumption $(B0)$ holds. Assume also that for all $k \in \mathcal{K}$, it holds that*

$$\lim_{T\to\infty}\sum_{t=1}^{T}\lambda_k(t) = +\infty \quad \text{and} \quad \lim_{T\to\infty}\sum_{t=1}^{T}\lambda_k(t)^2 < +\infty. \tag{4.37}$$

Then, if for all $k \in \mathcal{K}$, the SAA

$$\pi_k(n) \;=\; \pi_k(n-1) + \lambda_k(n)\left(\beta_k^{(\gamma_k)}\left(\hat{\mathbf{u}}_k(n)\right) - \pi_k(n-1)\right), \tag{4.38}$$

converges asymptotically to a vector $\pi^* = (\pi_1^*, \ldots, \pi_K^*)$ given the initial probability distributions $\pi(0) = (\pi_1(0), \ldots, \pi_K(0)) \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_1)$, it holds that, $\forall k \in \mathcal{K}$, $\pi_k^* = \beta_k\left(\bar{\mathbf{u}}_k\left(\cdot, \pi_{-k}^*\right)\right)$.

The proof of lemma 4.5.3 is presented in appendix K. Note also that Lemma 4.5.3 does not ensure the convergence. It rather states that if convergence is observed, then, the limiting probability distribution π_k^* represents a LE of the game $\mathcal{G}$ in SSI-BS. More specifically, from Prop. 4.5.1 it states that the behavioral strategy profile $\sigma \in \Sigma$ which generates the sequences $\{\beta_k^{(\gamma_k)}\left(\bar{\mathbf{u}}_k\left(\cdot, \pi_{-k}(n)\right)\right)\}_{n>0}$, is an ϵ-equilibrium of the game $\mathcal{G}$ if the resulting SAA given by (4.38) converges asymptotically. Note that the randomness in (4.38) comes from the fact that other players are also updating their own strategies. The conditions for observing convergence are analyzed later in Sec. 4.6.

4.5.3 Joint Learning of the Expected Performance and ϵ-equilibrium Strategies

Now, we drop both conditions $(A0)$ and $(B0)$ and we simultaneously analyze for all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, N_k\}$, the resulting coupled SAAs $\hat{u}_{k,A_k^{(n_k)}}(n)$ and

$\pi_{k,A_k^{(n_k)}}(n)$, with $n > 0$, i.e.,

$$\begin{cases} \hat{u}_{k,A_k^{(n_k)}}(n) &= \hat{u}_{k,A_k^{(n_k)}}(n-1)+ \\ & \alpha_k(n)\frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{\pi_{k,A_k^{(n_k)}}(n)}\left(\tilde{u}_{k(n-1)} - \hat{u}_{k,A_k^{(n_k)}(n-1)}\right), \\ \pi_{k,A_k^{(n_k)}}(n) &= \pi_{k,A_k^{(n_k)}}(n-1) \\ & +\lambda_k(n)\left(\beta^{(\gamma_k)}_{k,A_k^{(n_k)}}(\hat{\mathbf{u}}_k(n)) - \pi_{k,A_k^{(n_k)}}(n-1)\right). \end{cases} \tag{4.39}$$

In the following, we rely on the following assumption.

(B1) *For all $(j,k) \in \mathcal{K}^2$, the learning rates α_k and λ_j satisfy that*

$$\lim_{n\to\infty} \frac{\lambda_j(n)}{\alpha_k(n)} = 0. \tag{4.40}$$

Note that for all $(j,k) \in \mathcal{K}^2$ and for all $(n_j, n_k) \in \{1,\ldots,N_j\} \times \{1,\ldots,N_k\}$, the expected variation between two consecutive updates n and $n-1$ of the sthocastic approximation processes $\pi_{j,A_j^{(n_j)}}(n)$ and $\hat{u}_{k,A_k^{(n_k)}}(n)$ with respect to the learning rate α_k can be written as follows

$$\begin{aligned} \mathbb{E}_{\mathbf{h},\xi}\left[\left.\frac{\hat{u}_{k,A_k^{(n_k)}}(n) - \hat{u}_{k,A_k^{(n_k)}}(n-1)}{\alpha_k(n)}\right| \hat{\mathbf{u}}_k(n-1), \pi(n)\right] &= \\ \mathbb{E}_{\mathbf{h},\xi}\left[\left.\frac{\mathbb{1}_{\left\{a_k=A_k^{(n_k)}\right\}}}{\pi^{+}_{k,A_k^{(n_k)}}}\left(\tilde{u}_k - \hat{u}_{k,A_k^{(n_k)}}\right)\right| \hat{\mathbf{u}}_k(n-1), \pi(n)\right] &= \\ \bar{u}_k\left(\mathbf{e}_k^{(N_k)}, \pi_{-k}(n)\right) - \hat{\mathbf{u}}_{k,A_k^{(n_k)}}(n-1). \end{aligned}$$

and,

$$\begin{aligned} \mathbb{E}_{\mathbf{h},\xi}\left[\left.\frac{\pi_{j,A_j^{(n_j)}}(n) - \pi_{j,A_j^{(n_j)}}(n-1)}{\alpha_k(n)}\right| \pi_j(n-1), \hat{\mathbf{u}}_j(n)\right] &= \\ \left(\frac{\lambda_j(n)}{\alpha_k(n)}\right)\mathbb{E}_{\mathbf{h},\xi}\left[\left.\beta^{(\gamma_j)}_{j,A_j^{(n_j)}}(\hat{\mathbf{u}}_j(n)) - \pi_{j,A_j^{(n_k)}}(n-1)\right| \pi_j(n-1), \hat{\mathbf{u}}_j(n)\right], &= \\ \left(\frac{\lambda_j(n)}{\alpha_k(n)}\right)\left(\beta^{(\gamma_j)}_{j,A_j^{(n_j)}}(\hat{\mathbf{u}}_j(n)) - \pi_{j,A_j^{(n_k)}}(n-1)\right). \end{aligned}$$

respectively. From assumption $(B1)$, it can be stated that asymptotically, the SAA $\pi_j(n)$ can be seen as a time-invariant process by the performance-learning process $\hat{\mathbf{u}}_k(n)$. That is, as long as $n \longrightarrow \infty$, the expected variation of the SAA $\pi_j(n)$ with respect to the learning rate of the SAA $\hat{\mathbf{u}}_k(n)$ approaches zero. Following the same analysis, it can be also concluded that the SAA $\pi_k(n)$ sees the SAA $\hat{\mathbf{u}}_k(n)$ as a fast transient. Hence, it becomes natural to analyze the system of SAA in (4.39) considering the following hypothesis:

(i) For all $k \in \mathcal{K}$, the SAAs $\pi_1(n),\ldots,\pi_K(n)$ see the SAA $\hat{\mathbf{u}}_k(n)$ as always calibrated to the current vector $\pi(n) = (\pi_1(n),\ldots,\pi_K(n))$, for all $n \in \mathbb{N}$.

(ii) For all $k \in \mathcal{K}$, the SAAs $\hat{\mathbf{u}}_1(n), \ldots, \hat{\mathbf{u}}_K(n)$ see the SAA $\pi_k(n)$, as a time-invariant process.

These hypotheses imply that for all $k \in \mathcal{K}$, player k is able to learn their expected performance vector $\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n))$ before that all players change their corresponding probability distributions $\pi_k(n)$. The idea of using different learning rates, such that one SAA can be seen as a faster or slower than another one was originally introduced in [12] for the case of two coupled SAA. Few years after, it was extended to more than two coupled SAA in [48]. In this subsection, we follow the line of the work presented in [48].

The main result of this subsection is presented in the following theorem.

Theorem 4.5.4 *Consider the game $\mathcal{G}$ and assume that for all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, N_k\}$, it holds that for all $n \in \mathbb{N}$,*

$$\begin{cases} \hat{u}_{k,A_k^{(n_k)}}(n) &= \hat{u}_{k,A_k^{(n_k)}}(n-1) + \\ & \alpha_k(n) \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{\pi_{k,A_k^{(n_k)}}(n)} \left(\tilde{u}_{k}(n-1) - \hat{u}_{k,A_k^{(n_k)}}(n-1) \right), \\ \pi_{k,A_k^{(n_k)}}(n) &= \pi_{k,n_k}(n-1) + \\ & \lambda_k(n) \left(\beta_{k,n_k}^{(\gamma_k)}\left(\hat{\mathbf{u}}_k(n)\right) - \pi_{k,n_k}(n-1) \right), \end{cases} \tag{4.41}$$

where, $a_k(0) \in \mathcal{A}_k$, $\hat{\mathbf{u}}_k(0) \in \mathbb{R}^{N_k}$ and $\pi_k(0) \in \triangle(\mathcal{A}_k)$ are arbitrary initializations. Hence, under assumption (B1), if the set of SAAs (4.41) converge, it holds that,

$$\lim_{n\to\infty} \pi_k(n) = \pi_k^*, \tag{4.42}$$

$$\lim_{n\to\infty} \hat{u}_{k,A_k^{(n_k)}}(n) = \bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*), \tag{4.43}$$

where $\pi_k^* \in \triangle(\mathcal{A}_k)$ satisfies that,

$$\pi_k^* = \beta_k^{(\gamma_k)}\left(\bar{u}_k\left(\mathbf{e}_1^{(N_k)}, \pi_{-k}^*\right), \ldots, \bar{u}_k\left(\mathbf{e}_{N_k}^{(N_k)}, \pi_{-k}^*\right)\right). \tag{4.44}$$

The proof of Theorem 4.5.4 is an immediate result from Lemma 4.5.1 and Lemma 4.5.3. Prop. 4.5.1 states that the behavioral strategy profile $\sigma \in \Sigma$ which generates the sequences $\{\pi_k(n)\}_{n>0}$, is an ϵ-equilibrium of the game $\mathcal{G}$ if the resulting SAAs given by (4.41) converges asymptotically.

In the following section, we study the conditions over which convergence is observed.

4.6 Convergence Analysis

In this section, we study the convergence of the coupled learning dynamics presented in Theorem 4.5.4, under two particular scenarios. First, we consider that all players use comparable learning rates, i.e.,

(B2) *For any $(j,k) \in \mathcal{K}^2$ and $n > 0$, it holds that $\lambda_j(n) = b_{j,k}\lambda_k(n)$, where $b_{j,k} > 0$.*

Second, we consider that for all $k \in \mathcal{K} \setminus \{K\}$, the set $\mathcal{K} \setminus \{k\}$ can be portioned into two non-empty subsets. The first subset $\mathcal{K}_k^{(1)}$ contains players whose corresponding learning rate λ_j, with $j \in \mathcal{K}_k^{(1)}$, is slower that the learning rate λ_k, i.e., $\lambda_k > \lambda_j$ and thus, the SAA $\pi_j(n)$ can be considered as static with respect to the learning dynamics $\pi_k(n)$. The other set $\mathcal{K}_k^{(2)}$ contains players whose learning rate λ_j, with $j \in \mathcal{K}_k^{(2)}$ is faster that the learning rate λ_k, i.e., $\lambda_k < \lambda_j$ and thus, the SSA $\pi_j(n)$ can be considered as a fast transient with respect to the SAA $\pi_k(n)$. Without any loss of generality, we assume that $\mathcal{K}_k^{(1)} = \{1, \dots, k-1\}$ and $\mathcal{K}_k^{(2)} = \{k+1, \dots, K\}$. That is,

(B3) *For any $k \in \mathcal{K} \setminus \{K\}$, it holds that*

$$\lim_{n\to\infty} \frac{\lambda_k(n)}{\lambda_{k+1}(n)} = 0. \tag{4.45}$$

4.6.1 Scenario 1: Homogeneous Learning Rates

From the assumption $(B1)$, we have that, for all $k \in \mathcal{K}$, the SAAs $\hat{\mathbf{u}}_k(n)$ can be studied by considering all the SAAs $\pi_1(n), \dots, \pi_K(n)$ as time-invariant. Thus, from Lemma 4.5.1, it is known that for all $k \in \mathcal{K}$, the SAA $\hat{\mathbf{u}}_k(n)$ always converges to the limiting expected performance $\bar{\mathbf{u}}_k\left(\cdot, \pi^*_{-k}\right)$, where the vector π^* is the vector of limiting probability distributions to which the stochastic algorithms $\pi_1(n), \dots, \pi_K(n)$ converge to. Hence, under assumption $(B1)$, the convergence analysis of the coupled SAA in (4.41) reduces to the analysis of convergence of the SAAs $\pi_1(n), \dots, \pi_K(n)$. For doing such an analysis, we rely on the ordinary differential equation (ODE) approximations obtained in the proof of Lemma 4.5.3, i.e.,

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_1(t) & = & \beta_1^{(\gamma_1)}\left(\bar{\mathbf{u}}_1\left(\cdot, \widehat{\pi}_{-1}(t)\right)\right) - \widehat{\pi}_1(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_2(t) & = & b_{2,1}\left(\beta_2^{(\gamma_2)}\left(\bar{\mathbf{u}}_2\left(\cdot, \widehat{\pi}_{-2}(t)\right)\right) - \widehat{\pi}_2(t)\right), \\ & \vdots & \\ \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_K(t) & = & b_{K,1}\left(\beta_K^{(\gamma_K)}\left(\bar{\mathbf{u}}_K\left(\cdot, \widehat{\pi}_{-K}(t)\right)\right) - \widehat{\pi}_K(t)\right). \end{cases} \tag{4.46}$$

The set of ODEs in (4.46) turns out to be the same set of ODEs describing the smooth fictitious play [9]. Such a system of differential equations has been studied in [37] and therein, it has been shown that there exists a Lyapunov function for two-player zero-sum games, which implies that the trajectories described by (4.46) are globally convergent. Using the same reasoning, potential games also possess a Lyapunov function, which is indeed, the potential function of the game [84]. Hence, the following proposition holds.

Theorem 4.6.1 (Sufficient Conditions for Convergence) *Consider the game $\mathcal{G}$ and assume that assumptions $(B1)$ and $(B2)$ hold. Then, the coupled SAAs in Theorem 4.5.4 converge if the game $\mathcal{G}$ belongs to one of the following classes,*

- *Potential games,*
- *Zero-sum game with $K = 2$ players.*

- *Dominance solvable games,*
- *Games with unique evolutionary stable strategy.*

It is important to remark that many of the problems found in wireless communications are potential games [64], thus the behavioral strategies described in the previous section are naturally ϵ-equilibrium behavioral strategies in this class of games. The interest reader is referred to [10, 73, 80, 96].

4.6.2 Scenario 2: Heterogeneous Learning Rates

Following the same reasoning of the previous section, under the assumption $(B1)$ and $(B3)$, it becomes natural to analyze the coupled SAAs in Theorem 4.5.4 considering the hypotheses (i) and (ii) described in Sec. 4.5.3 and the following hypothesis:
(iii) For all $k \in \mathcal{K}$, the SAA $\pi_k(n)$ sees the SAA $\pi_j(n)$, with $j \in \mathcal{K} \setminus \{k+1, \ldots, K\}$, as time-invariant processes.
(iv) For all $k \in \mathcal{K}$, the SAA $\pi_k(n)$ sees the SAA $\pi_j(n)$, with $j \in \mathcal{K} \setminus \{1, \ldots, k-1\}$, as always calibrated to the current values $\pi_1(n), \ldots, \pi_k(n)$.
Under hypothesis (iii), we can analyze the SAA $\pi_K(n)$, $n > 0$, assuming that players $1, \ldots, K-1$ play a given SSI-BS profile represented by the vector $\pi^{+}_{-K} = \left(\pi^{+}_{1}, \ldots, \pi^{+}_{K-1}\right) \in \triangle\left(\mathcal{A}_1\right) \times \ldots \times \triangle\left(\mathcal{A}_{K-1}\right)$. For any mixed strategy π^{+}_{-K}, the SSA $\pi_K(n)$ in Theorem 4.5.4 asymptotically approximates the solution of the following ODE,

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_K(t) \quad = \quad \beta_K^{(\gamma_K)}\left(\bar{\mathbf{u}}_K\left(\cdot, \pi_1^{+}, \ldots, \pi_{K-1}^{+}\right)\right) - \widehat{\pi}_K(t). \tag{4.47}$$

Thus, independently of the initial probability distribution $\pi_k(0)$, the unique asymptotically stable point of the ODE in (4.47) is

$$\pi_K = \beta_K^{(\gamma_K)}\left(\bar{\mathbf{u}}_K\left(\cdot, \pi_1^{+}, \ldots, \pi_{K-1}^{+}\right)\right). \tag{4.48}$$

Following the hypotheses (iii) and (iv), the SAA $\pi_{K-1}(n)$ can be analyzed using the continuous time process $\widehat{\pi}_{K-1}(t)$, $t \in [0, \infty[$, which approximates the ODE,

$$\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_{K-1}(t) = \\ \beta_{K-1}^{(\gamma_{K-1})}\left(\bar{\mathbf{u}}_{K-1}\left(\cdot, \pi_1^{+}, \ldots, \pi_{K-2}^{+}, \beta_K^{(\gamma_K)}\left(\bar{\mathbf{u}}_K\left(\cdot, \pi_1^{+}, \ldots, \pi_{K-2}^{+}, \widehat{\pi}_{K-1}(t)\right)\right)\right)\right) \\ -\widehat{\pi}_{K-1}(t). \end{aligned} \tag{4.49}$$

Note that if the trajectory described by the ODE (4.49) converges to a given stable equilibrium denoted by $b_{K-1}\left(\pi_1, \ldots, \pi_{K-2}\right)$, where the mapping

$$b_{K-1} : \triangle(\mathcal{A}_1)\times, \ldots, \times\triangle(\mathcal{A}_{K-2}) \to \triangle(\mathcal{A}_{K-1}) \tag{4.50}$$

is Lipchitz, then the SAA $\pi_{K-2}(n)$ can be studied following the same methodology used for the SAA $\pi_K(n)$, i.e., ODE approximation. Hence, in order to study all the following strategy-learning processes $\pi_j(n)$, for all $j \in \{2, \ldots, K-1\}$, the following general assumption is stated.

(B4) *For all $j \in \{2, \ldots, K-1\}$, there exists a mapping $b_j : \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_{j-1}) \rightarrow \triangle(\mathcal{A}_j)$ which is Lipschitz and $b_j\left(\pi_1^+, \ldots, \pi_{j-1}^+\right)$ is the globally asymptotically stable equilibrium point of the ODE*

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_j(t) \quad = \quad \beta_j^{(\gamma_j)}\left(\bar{\mathbf{u}}_j\left(\cdot, \pi_1^+, \ldots, \pi_{j-1}^+, \mathbf{B}_j\right)\right) - \widehat{\pi}_j(t), \tag{4.51}$$

where the vector $\mathbf{B}_j = (B_{j+1}, \ldots, B_K) \in \triangle(\mathcal{A}_{j+1}) \times \ldots \times \triangle(\mathcal{A}_K)$ is defined as follows,

$$B_{j+1} \quad = \quad b_{j+1}\left(\pi_1^+, \ldots, \pi_{j-1}^+, \widehat{\pi}_j(t)\right) \tag{4.52}$$

and $\forall m \in \{2, \ldots, K-j-1\}$,

$$B_{j+m} \quad = \quad b_{j+m}\left(\pi_1^+, \ldots, \pi_{j-1}^+, \widehat{\pi}_j(t), B_{j+1}, \ldots, B_{j+m-1}\right)$$

and

$$B_K \quad = \quad \beta_K^{(\gamma_K)}\left(\pi_1^+, \ldots, \pi_{j-1}^+, \widehat{\pi}_j(t), B_{j+1}, \ldots, B_{K-1}\right).$$

Note that if the assumption **(B4)** holds, the analysis of the coupled strategy learning in (4.36) reduces to the study of the following ODE,

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_1(t) \quad = \quad \beta_1^{(\gamma_1)}\left(\bar{\mathbf{u}}_1\left(\cdot, \mathbf{B}_1\left(\widehat{\pi}_1(t)\right)\right)\right) - \widehat{\pi}_1(t). \tag{4.53}$$

If the trajectory described by the ODE (4.53) converges to a given vector $\pi_1^* \in \triangle(\mathcal{A}_1)$, it follows from (4.51) that

$$\pi_1^* = \beta_1\left(\bar{\mathbf{u}}_1\left(\cdot, \mathbf{B}_1\left(\pi_1^*\right)\right)\right). \tag{4.54}$$

Moreover, from assumption **(B4)**, it holds that $\forall j \in \{2, \ldots, K-1\}$,

$$\pi_j^* \quad = \quad \beta_j^{(\gamma_j)}\left(\bar{\mathbf{u}}_j\left(\cdot, \pi_1^*, \ldots, \pi_{j-1}^*, \mathbf{B}_j\left(\pi_1^*, \ldots, \pi_j^*\right)\right)\right),$$

which together with (4.48) prove the following Theorem.

Theorem 4.6.2 (Sufficient Conditions for Convergence) *Consider the game $\mathcal{G}$ and assume that assumptions ($B1$), ($B3$) and ($B4$) hold. Then, the coupled SSAs in Theorem 4.5.4 converges if the following ODE,*

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_1(t) \quad = \quad \beta_j\left(\bar{\mathbf{u}}_j\left(\cdot, \mathbf{B}_1\left(\widehat{\pi}_1(t)\right)\right)\right) - \widehat{\pi}_1(t) \tag{4.55}$$

has a stable solution.

From theorem 4.6.2, it can be implied that to show the convergence of the coupled SAAs in Theorem 4.5.4, it suffices to show that conditions ($B1$), ($B3$) and ($B4$) are satisfied and to analyze the convergence of the SSA $\pi_1(n)$ by following the ODE (4.55), i.e., the slowest learner. It is important to note that in [48], it has been shown that the coupled SAAs $\pi_1(n), \ldots, \pi_K(n)$ are described by the system of differential equations given by (4.41) when the coefficients $c_{j,1}$, for all $j \in \mathcal{K}$ satisfies that $\forall j \in \{1, \ldots, K-1\}$,

$$c_{j+1,1} = o\left(c_{j,1}\right) \text{ as } c_{j,1} \to 0. \tag{4.56}$$

This result implies that the coupled SAA in (4.41) under conditions ($B1$) and ($B3$) inherits the same convergence properties that the SAA in (4.41) under conditions ($B1$) and ($B2$). However, there exist two classes of games known as Jordan's matching pennies game [39] and Shapley's variant of rock-scissors-paper game [103] where the coupled SAA in (4.41), following assumptions ($B1$) and ($B2$), does not converge. On the contrary, following assumptions ($B1$) and($B3$) convergence is observed [48]. In the following section, we provide another example where convergence depend on the assumptions over the learning rates, i.e., the pairs ($B1$,$B2$) or ($B1$, $B3$).

4.7 Applications

In this section, we provide a numerical analysis of the performance achieved by radio devices following the behavioral rule proposed in this chapter. First, we focus on the achieved performance in limited time. Here, our interest focuses in determining the sum spectral efficiency when the learning period is limited. Note that the theoretical analysis requires infinite time for convergence, which is not practically appealing. Second, we focus on the impact of the number of choices each radio devices might possess at a given time. Here, we verify the counter intuitive result which states that increasing the set of choices each radio device possesses during the whole game realization might lead to worse global performance. In the following, we describe the scenarios used to highlight these findings.

4.7.1 Logit Equilibrium and Interference Channels

Consider a set $\mathcal{K} = \{1, \ldots, K\}$ of transmitter-receiver pairs. Each transmitter sends private information to its respective receiver trough a set $\mathcal{S} \triangleq \{1, \ldots, S\}$ of orthogonal channels. Here, the orthogonality is assumed in the frequency domain. All transmitters simultaneously use the same set $\mathcal{S}$ of channels and thus, communications are subject to mutual interference. Let $h_{j,k}^{(s)}(n)$ represent the channel realization between transmitter k and receiver j over channel s at time n. In our analysis, flat fading channels are assumed during the symbol period, i.e., the channel realization is assumed time-invariant during the transmission of one symbol, however, the channel might vary from symbol to symbol period. Denote by $\mathbf{h}(n) = \left(h_{j,k}^{(s)}(n)\right) \in \mathbb{C}^{J \cdot K \cdot S}$ the vector of channel realizations at interval n and let $\mathcal{H}$ be the finite set of all possible channel realization vectors (in practice, relevant quantities like channel quality indicators in 3G cellular systems are quantized). Let $\mathbf{h}^{(i)}$ be the i-th element of the set $\mathcal{H}$, with $i \in \{1, \ldots, |\mathcal{H}|\}$. For each channel use, the vector $\mathbf{h}(n)$ is drawn from the set $\mathcal{H}$ following a probability distribution $\rho = (\rho_{\mathbf{h}^{(1)}}, \ldots, \rho_{\mathbf{h}^{(|\mathcal{H}|)}}) \in \triangle(\mathcal{H})$. That is, $\rho_{\mathbf{h}^{(i)}} = \Pr\left(\mathbf{h}(n) = \mathbf{h}^{(i)}\right)$, for all $n \in \mathbb{N}$. The vector of transmitted symbols

$\mathbf{x}_k(n)$ is an S-dimensional random variable with zero mean and covariance matrix $\mathbf{P}_k(n) = \mathbb{E}\left(\mathbf{x}_k(n)\mathbf{x}_k^H(n)\right) = \operatorname{diag}\left(p_{k,1}(n), \ldots, p_{k,S}(n)\right)$. For all $(k,s) \in \mathcal{K} \times \mathcal{S}$, $p_{k,s}(n)$ represents the transmit power allocated by transmitter k over channel s. A power allocation (PA) vector for transmitter $k \in \mathcal{K}$ is any vector

$$\mathbf{p}_k(n) = (p_{k,1}(n), \ldots, p_{k,S}(n)) \in \mathcal{A}_k,$$

where,

$$\begin{aligned} \mathcal{A}_k \quad = \quad & \left\{\mathbf{p}_k^{(s)} = p_{k,\max}\,\mathbf{e}_s : \forall s \in \mathcal{S},\ \mathbf{e}_s = (e_{s,1}, \ldots, e_{s,S}),\right. \\ & \left. \forall r \in \mathcal{S} \setminus s,\ e_{s,r} = 0, \text{ and } e_{s,s} = 1\right\}. \end{aligned} \tag{4.57}$$

and $p_{k,\max}$ is the maximum transmit power of transmitter k. Following this notation, the power allocation vector $\mathbf{p}_k^{(s)}$ represents the s-th element of the set $\mathcal{A}_k$. We denote by $N_k = |\mathcal{A}_k|$ the cardinality of the set $\mathcal{A}_k$. We respectively denote the noise spectral density and the bandwidth of channel $s \in \mathcal{S}$ by N_0 and B_s. The total bandwidth is denoted by $B = \sum_{s=1}^{S} B_s$, independently of the receiver. We denote the individual spectral efficiency of transmitter $k \in \mathcal{K}$ as follows,

$$u_k(\mathbf{h}(n), \mathbf{p}_k(n), \mathbf{p}_{-k}(n)) = \sum_{s \in \mathcal{S}} \frac{B_s}{B} \log_2\left(1 + \gamma_{k,s}(n)\right) \text{ [bps/Hz]}, \tag{4.58}$$

where $\gamma_{k,s}(n)$ is the signal-to-interference plus noise ratio (SINR) seen by player k over its channel s at time n, i.e.,

$$\gamma_{k,s}(n) = \frac{p_{k,s}(n) g_{j,k}^{(s)}(n)}{N_0 B_s + \sum\limits_{i \in \mathcal{K} \setminus \{k\}} p_{i,s}(n) g_{j,i}^{(s)}(n)}. \tag{4.59}$$

Here, for all $(j,k) \in \mathcal{K}^2$ and $n \in \mathbb{N}$, $g_{j,k}^{(s)}(n) \triangleq \left|h_{j,k}^{(s)}(n)\right|^2$.

4.7.2 Convergence in Finite Time

In order to run a fair comparison of the behavioral rule in Theorem 4.5.4 with existing results, e.g., best response dynamics, fictitious play, regret matching learning, and cumulative payoff matching reinforcement learning (CPM-RL), we consider the set $\mathcal{H}$ is unitary. This implies that the stochastic game reduces to play the same one-shot game repeatedly *ad infinitum*. At each stage, the corresponding one-shot game might have either one NE in pure strategies or two NE in pure strategies plus one NE in mixed strategies [90], depending on the particular channel realization. Here, we generate $10,000$ channel realizations. For each channel realization, we calculate the sum of individual spectral efficiencies at the NE, using the theoretical results in [90]. We consider the average network spectral efficiency (NSE) at the best NE and the worst NE. That is, the average NSE at the NE with the highest NSE and lowest NSE, respectively. Similarly, for each channel realization, we determine the sum of individual spectral efficiencies achieved by both transmitters using the BRD, FR,

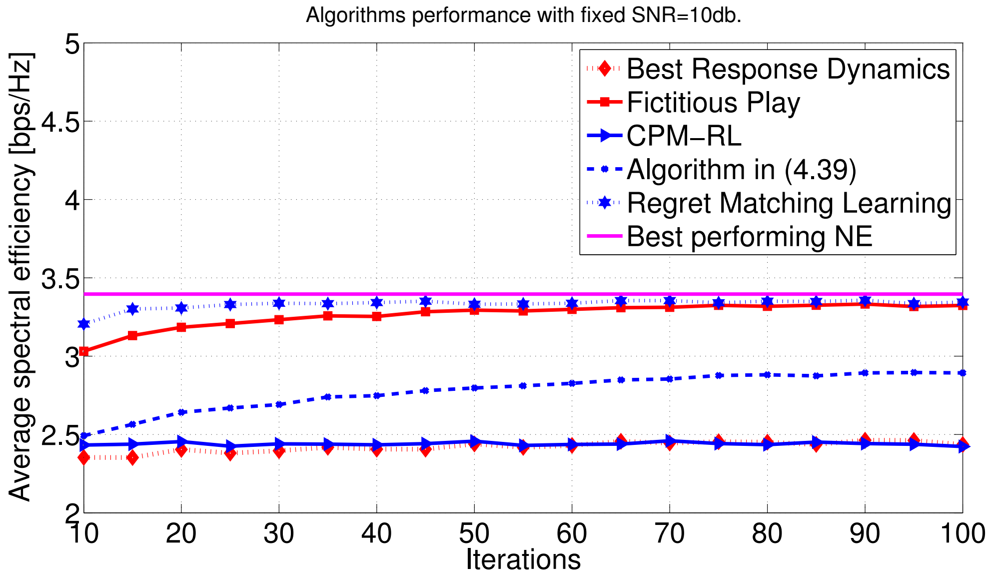

Figure 4.1: Average sum spectral efficiency in the two-transmitter two-receiver two-channel game, when transmitters are limited to channel selection. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{(\frac{3}{4})}}$, $\alpha_2(n) = \frac{1}{n^{(\frac{2}{3})}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs.

CPM-RL and the proposed behavioral rule in Theorem 4.5.4, when their learning time is limited to a fixed number of time intervals (game repetitions).

In Fig. 4.1, we plot the average sum of individual spectral efficiencies as a function of the number of time intervals the players are let to interact. Naturally, the algorithm in Theorem 4.5.4 performs better with higher number of iterations. However, an important observation here is that it achieves a higher performance than the classical (simultaneous) best response dynamics (Def. 3.3.4) and cumulative payoff matching reinforcement learning (CPM-RL) [120]. This confirms the intuition presented in this chapter that BRD and reinforcement learning, in general, converge to action profiles which might not be an equilibrium, or even worst, action profiles which are suboptimal from an individual and global point of view. Particular attention must be put to the fact that in the case of the algorithm in Theorem 4.5.4 and CPM-RL, both algorithms require the same amount of information. However, the algorithm presented in Theorem 4.5.4 performs better.

Another important remark in Fig. 4.1 is the fact that, fictitious play (See Sec. 3.4.4) and regret matching learning [35,90] outperform the algorithm proposed in Theorem 4.5.4. However, those algorithms require both the observation of the actions of the other player and the knowledge of the explicit expression of the utility function at each game stage.

In Fig. 4.2, we plot the average sum spectral efficiency as a function of the SNR, when the number of iterations has been fixed to 40 iterations. Note that in terms of performance, the behavior is the same described in the previous figure. However, an important remark here is the fact that at high SNR, the game possesses two NE in pure strategies [90]. Thus, the increasing gap between the NE with the highest

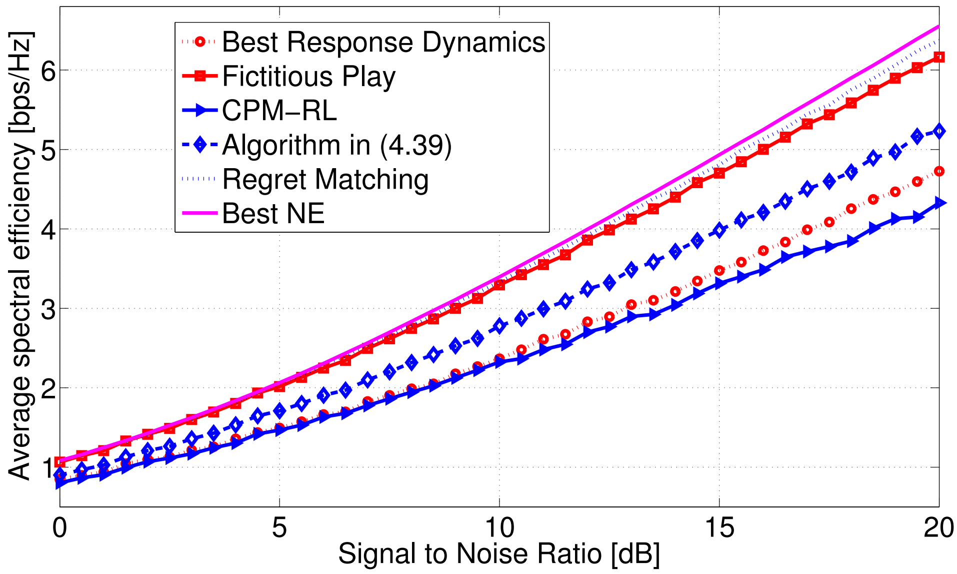

Figure 4.2: Average sum spectral efficiency in the two-transmitter two-receiver two-channel game, when transmitters are limited to channel selection. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{\left(\frac{3}{4}\right)}}$, $\alpha_2(n) = \frac{1}{n^{\left(\frac{2}{3}\right)}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs.

average performance and the actually achieved performance is due to the fact that the algorithm in might converge to any of both equilibria, i.e., not always to the best NE. In the case of BRD and CPM-RL, they can converge to action profiles which perform worse than the worst NE. Hence, their average performance is always worst that the algorithm presented in (4.39). On the contrary, regret matching learning and FP seems to converge always to the best NE. However, it does not exist a formal proof of this observation.

4.7.3 Impact of the Number of Choices

In this subsection, we increase the number of available channels and we let each transmitter to use either a unique channel or any subset of adjacent channels. Thus, if we consider S channels, the cardinality of the sets $\mathcal{A}_k$ is $\frac{S}{2}(1+S)$, for all $k \in \mathcal{K}$. Here, we generate 10000 channel realizations and we let radio devices to learn through 100 game repetitions. In Fig. 4.3, it is shown that increasing the number of available channels leads to a loss of spectral efficiency. This observation is due to the fact that letting each radio device to use an additional channel implies increasing the number of actions. Then, the radio devices must spend more time on testing all their actions to build the vector of utility estimations rather that using the optimal action. This effect, significantly reduces the average individual spectral efficiency. However, the converging point is still a logit equilibrium (Def. 4.4.2).

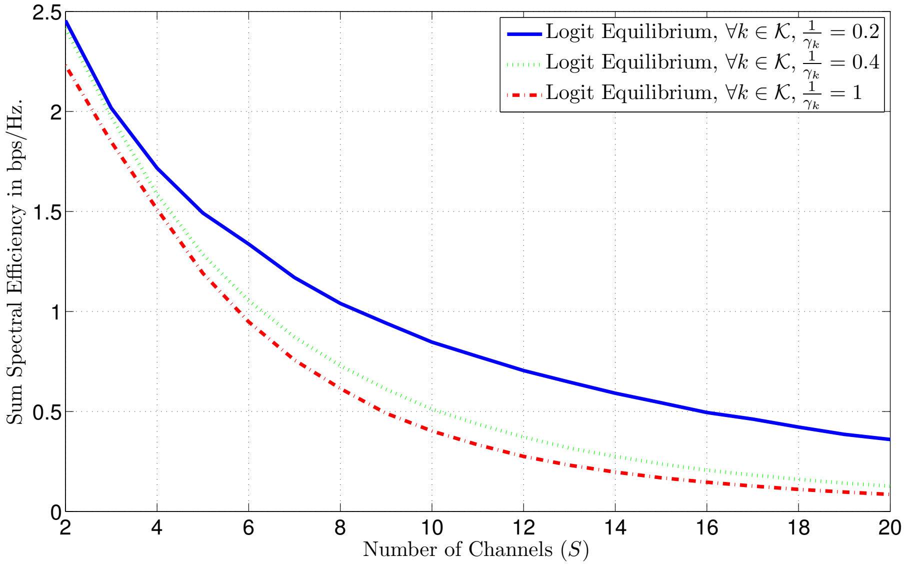

Figure 4.3: Achieved sum spectral efficiency in the two-transmitter two-receiver N-channel game, when transmitters are limited to use only one channel or any combination of adjacent channels. Here $\alpha_1(n) = \alpha_2(n) = \frac{1}{n^{(\frac{3}{4})}}$, $\alpha_2(n) = \frac{1}{n^{(\frac{2}{3})}}$ and $\lambda_1(n) = \frac{1}{n}$. Moreover, $SNR = \frac{p_{k,\max}}{\sigma^2} = 10$ dBs.

4.8 Conclusions

In this chapter, learning dynamics adapted to real-system implementation constraints have been introduced to learn equilibrium in fully decentralized wireless networks. For instance, it has been assumed that the only information a radio device can obtain from the network is a measure of its instantaneous performance, and, each radio device is completely unaware of the existence of all the other radio devices. Under these conditions, we introduced novel learning dynamics such that each radio device is able to simultaneously learn both the optimal mixed strategy at the equilibrium and the expected performance achieved with each of its actions. Here, we have used recent tools from stochastic approximations to study the convergence of such dynamics. In particular, we show that there exist several classes of games where these learning dynamics always converge to LE. Using a numerical analysis, we show that the proposed dynamics perform better than cumulative payoff matching reinforcement learning, which requires the same information knowledge. It also outperforms the classical BRD, which is in general more demanding in terms of information. We also observed that the proposed technique is outperformed by classical FP and regret matching learning. However, both techniques require the knowledge of a closed form expression of the utility function and the observation of the actions of all the other players at each learning stage.

Chapter 5

QoS Provisioning in Spectrum Sharing Games

In this chapter, we study an equilibrium concept, namely the satisfaction equilibrium (SE), where in contrast to existing equilibrium notions, for instance Nash equilibrium (NE) and generalized NE (GNE), the idea of performance optimization in the sense of utility maximization or cost minimization does not exist. The concept of SE relies on the fact that players might be either satisfied or unsatisfied with their achieved performance. At the SE, if it exists, all players are satisfied. This notion of equilibrium perfectly models the problem of QoS provisioning in unmanaged spectrum access. Here, radio devices are satisfied if they are able to provide the requested QoS.

In the context of unmanaged spectrum access, many equilibrium concepts have been studied and proved to be interesting game outcomes [91], e.g., coarse correlated equilibrium (CCE) [120], correlated equilibrium (CE) [3], Nash equilibrium (NE) [66] and generalized NE (GNE) [24]. For instance, the notion of NE has been used to determine individually optimal power allocation policies in decentralized self-configuring networks where radio devices are interested in maximizing their individual transmission rates [27,80,97,99,100], energy efficiency [60,61] , or minimizing the outage probability [7]. The idea of GNE has been used in [71] to minimize the transmit power consumption while guaranteeing some individual minimum signal to interference plus noise ratios (SINR). The concept of correlated equilibrium has been used in [1] to analyze the multiple access in wireless networks.

In particular, all the equilibrium notions mentioned above rely on the idea of individual utility maximization or individual payoff minimization. The difference between those equilibrium concepts depends mostly in the degree of correlation between the actions of the players (See [120] and references therein). However, from a practical point of view, a network operator, service provider or even the final user might be more interested in profiting from a minimum performance required to communicate (Quality of Service (QoS) requirement) rather than attaining the highest achievable performance. For instance, if a voice message is to be transmitted, a typical utility function (assuming the network latency is shorter than 200 ms) for such a device is the transmission rate. In any case, once the radio device has achieved a transmission configuration such that it is possible to ensure a minimum of 8000 samples/sec [62], there is no particular interest from such a device in changing its own transmission configuration. This is basically because the voice sampling rate is fixed and beyond a given threshold, human ears cannot detect any difference. The same reasoning can be applied to most of the communications services (with its corresponding QoS metric), e.g., data and video streaming, considering latency, transmission rate, etc. Thus, this might imply that, in practical terms, the classical game formulation with utility maximization or cost minimization might not properly model the QoS provisioning problem in self-configuring wireless communications.

In this chapter, we present a notion of equilibrium, namely satisfaction equilibrium (SE), which models more precisely the effect that in decentralized self-configuring networks, a communication takes place successfully if certain (minimum QoS) conditions are guaranteed. The notion behind the SE relies on the fact that players are either satisfied or unsatisfied with their actual performance. One player is said to be satisfied, if it plays an action which satisfies certain conditions given the actions of all the other players. A game is said to be in SE, if the game possesses one, when all players are satisfied. The idea behind SE was originally introduced in [93, 94] for a particular class of conditions in pure strategies. Later, the concept was formulated in terms of a fixed point inclusion for the case of pure strategies in [83,85]. As stated in [85], the advantages of the notion of SE over the classical notions such as NE and GNE, at least in the domain of signal processing for wireless communications are manifold. Here, we highlight the fact that, (*i*) the existence of the SE is less restrictive than the notion of GNE. That is, the network can possess an equilibrium in the sense of satisfaction but not in the sense of utility maximization. (*ii*) The behavioral rules needed for radio devices to learn a SE are by far simpler that behavioral rules to learn NE or GNE [91]. As we shall see, the only information required by each radio device at each time slot is whether it is satisfied or not (1-bit message). (*iii*) Using particular behavioral rules, convergence to SE can be achieved in finite time. This drastically contrast with the behavioral rules required to learn other equilibrium concepts. Here, at least a numerical value (several bits) of the instantaneous achieved utility is required and convergence is observed asymptotically. Within this framework, the contributions presented in this chapter are the followings:

- The notion of SE is formalized as a fixed point inclusion for both pure strategies (PS) and mixed strategies (MS). Conditions for the existence of the SE in PS and MS are established. Finally, the notion of SE is compared with existing equilibrium concepts, mainly, the NE [66] and generalized NE (GNE) [24].

- We introduce the notion of epsilon-satisfaction equilibrium (ϵ-SE), which consists of a mixed strategy which allows all players to be satisfied with probability not less than a given positive epsilon. This equilibrium concept turns out to be less restrictive in terms of existence. However, it is shown that not every game possesses an ϵ-SE.

- A refinement of the notion of SE to which we refer as efficient SE (ESE) is presented. The ESE relies on the idea of effort or cost of satisfaction. Here, each player independently ranks its own actions in terms of effort. At the ESE, if it exists, all players achieve satisfaction by using the transmit configuration which requires the lowest effort. Sufficient conditions for the existence and uniqueness of the ESE are also presented.

- A behavioral rule that allows radio devices to achieve the SE in PS, when it exists, is presented. Interestingly, following this rule, the convergence to the SE is observed in finite time and it requires only 1-bit feedback between the corresponding transmitter-receiver pairs at each game stage.

The sequel of the chapter is organized as follows. In Sec. 5.1, the QoS provisioning problem in decentralized self-configuring networks is formulated. In Sec. 5.2, a novel game formulation, called satisfaction form, is introduced as well as its extension in mixed strategies. The concept of SE and ϵ-SE is presented in the context of games in satisfaction form. Therein, a simple example is used to evidence the relevance of this concepts the context of QoS provisioning. In Sec. 5.3, the existence and uniqueness of the SE is analyzed. In Sec. 5.5, a refinement of the SE, which we call efficient SE, is introduced. In Sec. 5.4, we compare the notion of SE with existing equilibrium notions, such as NE and GNE in the context of the QoS provisioning problem. In Sec. 5.6, behavioral rules that allow radio devices to learn a SE are described. In Sec. 5.7, the notion of SE is used in the context of the interference channel where transmitters must guarantee a minimum transmission rate. Therein, it is shown that contrary to the notions of SE and ESE, classical equilibrium concepts fail to simultaneously satisfy both transmitters. The chapter is concluded by Sec. 5.8.

5.1 Problem Formulation

In general, the term QoS provisioning refers to all the procedures carried out by the radio devices aiming to guarantee a satisfying final user communication experience. That is, guaranteeing that communications take place with acceptable data rates, frame delays, spectral efficiency, energy efficiency, etc. In order to achieve such a goal, radio devices adjust their transmit and/or receive configurations. For instance, transmitters might tune parameters such as their channel selection and power allocation policy, modulation and error correction schemes, constellation sizes, etc. Similarly, receivers might tune their scheduling, decoding order, etc. The key point in this transmit/receive configuration tuning is that any change of a particular radio device on its own configuration influences the performance (QoS) of other radio devices. In the physical layer, this is basically, due to the mutual interference.

In the particular case of decentralized self-configuring networks, message exchanging between radio devices for establishing a sort of coordination to jointly improve the individual or global performance is highly constrained. This is basically because of the amount of signaling it might require and also because of the use of different physical layer technologies. Within this framework, it is common to model the radio devices as selfish entities concerned only with their own individual performance. In the sequel of this chapter, we study a particular QoS policy where all radio devices aim to satisfy a given QoS condition, for instance, a minimum data rate and packet delay. Here, none of the radio devices is interested in maximizing any of its performance metrics. Our goal then, is to provide a mathematical framework for the study of this scenario and designing the behavioral strategies that allow, if possible, the QoS satisfaction of all the radio devices.

5.2 Games in Satisfaction Form and Satisfaction Equilibrium

In this section, we introduce a novel game formulation where in contrast to existing formulations (e.g., normal form [66] and normal form with constrained action sets [24]), the idea of performance optimization, i.e., utility maximization or cost minimization, does not exist. In our formulation, to which we refer as satisfaction-form, the aim of the players is to adopt any of the actions which allows them to satisfy a specific condition given the actions adopted by all the other players. Under this game formulation, we introduce the concept of satisfaction equilibrium.

5.2.1 Games in Satisfaction Form

In general, a game in satisfaction-form can be described by the following triplet

$$\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}. \tag{5.1}$$

Here, the set $\mathcal{K} = \{1, \ldots, K\}$ represents the set of players and the set $\mathcal{A}_k = \{A_k^{(1)}, \ldots, A_k^{(N_k)}\}$ represents the set of N_k actions available for transmitter k. An action profile is a vector $\mathbf{a} = (a_1, \ldots, a_K) \in \mathcal{A}$, where,

$$\mathcal{A} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_K. \tag{5.2}$$

In this analysis, the set $\mathcal{K}$ is assumed finite, non-empty and non-unitary. The set $\mathcal{A}$ can be either finite or compact and convex. When necessary, the distinction is done explicitly. We denote by $\mathbf{a}_{-k} = (a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_K) \in \mathcal{A}_{-k}$, where

$$\mathcal{A}_{-k} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_{k-1} \times \mathcal{A}_{k+1} \times \ldots, \times\mathcal{A}_K, \tag{5.3}$$

the vector obtained by dropping off the k-th component of the vector $\mathbf{a}$. With a slight abuse of notation, we write the vector $\mathbf{a}$ as $(a_k, \mathbf{a}_{-k})$, in order to emphasize its k-th component. The correspondence $f_k : \mathcal{A}_{-k} \rightarrow 2^{\mathcal{A}_k}$ determines the set of actions of player k which allows its satisfaction given the actions played by all the other players. Here, the notation $2^{\mathcal{A}_k}$ refers to the set of all possible subsets of the set $\mathcal{A}_k$, including $\mathcal{A}_k$.
In the following example, we show how a given decentralized self-configuring network can be modeled by a game in satisfaction form.

Example 5.2.1 *Consider a decentralized and self-configuring network made of a set $\mathcal{K} = \{1, \ldots, K\}$ of transmitter-receiver pairs. For all $k \in \mathcal{K}$, let $\mathcal{A}_k$ be the* *set of transmit configurations available for transmitter k and let the function $u_k : \mathcal{A}_1 \times \ldots \times \mathcal{A}_K \to \mathbb{R}$ denote its (Shannon) transmission rate. Transmitter k must guarantee a data rate higher than Γ_k bps. Hence, the set of configurations it must adopt, given the configurations $\mathbf{a}_{-k}$ of all the other transmitters, is determined by the correspondence $f_k : \mathcal{A}_{-k} \to 2^{\mathcal{A}_k}$, which we define as follows,*

$$f_k\left(\mathbf{a}_{-k}\right) = \left\{a_k \in \mathcal{A}_k : u_k\left(a_k, \mathbf{a}_{-k}\right) \geqslant \Gamma_k\right\}. \tag{5.4}$$

Thus, the behavior of this network can be modeled by the game in satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$.

In Sec. 5.4, we use this example to show the differences between the satisfaction form and the normal form.
In general, an important outcome of a game in satisfaction form is the one where all players are satisfied. We refer to this game outcome as a satisfaction equilibrium (SE).

Definition 5.2.1 (Satisfaction Equilibrium in PS [85]) *An action profile $\mathbf{a}^+$ is an equilibrium for the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$ if*

$$\forall k \in \mathcal{K}, \quad a_k^+ \in f_k\left(\mathbf{a}_{-k}^+\right). \tag{5.5}$$

Note that in this formulation, radio devices are indifferent to the fact that there might exist another transmit configuration with which a higher utility e.g., transmission rate, can be obtained. Here, as long as each player is able to satisfy its individual conditions, it has no incentive to deviate from the current action profile. In the following, we describe the extension in mixed strategies of the game in satisfaction form.

5.2.2 Extension in Mixed Strategies of the Satisfaction Form

The concept of mixed strategies was introduced by Borel in [26]. A mixed strategy of player k is a probability distribution over the set of actions $\mathcal{A}_k$. We denote the set of all possible probability distributions over the set $\mathcal{A}_k$ by $\triangle\left(\mathcal{A}_k\right)$, i.e., the unit simplex over the elements of $\mathcal{A}_k$. We denote by $\pi_k = \left(\pi_{k,A_k^{(1)}}, \ldots, \pi_{k,A_k^{(N_k)}}\right)$ the probability distribution (mixed strategy) chosen by player k. Here, for all $n_k \in \{1, \ldots, N_k\}$, $\pi_{k,A_k^{(n_k)}}$ represents the probability that player k plays action $A_k^{(n_k)} \in \mathcal{A}_k$.
Following this notation, we denote by $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \left\{\triangle\left(\mathcal{A}_k\right)\right\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$ the extension in mixed strategies of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$, where the correspondence

$$\bar{f}_k : \prod_{j\in\mathcal{K}\setminus\{k\}} \triangle\left(\mathcal{A}_j\right) \to 2^{\triangle\left(\mathcal{A}_k\right)}, \tag{5.6}$$

determines the set of all possible probability distributions that allow player k to always choose an action which satisfies its individual conditions, that is,

$$\bar{f}_k\left(\pi_{-k}\right) = \left\{\pi_k \in \triangle\left(\mathcal{A}_k\right) : \Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\left(\pi_k, \pi_{-k}\right) = 1\right\}. \tag{5.7}$$

In this context, we define the SE as follows.

Definition 5.2.2 (Satisfaction Equilibrium in MS) *The mixed strategy profile $\pi^* \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ is a SE of the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$, if for all $k \in \mathcal{K}$,*

$$\pi_k^* \in \bar{f}_k\left(\pi_{-k}^*\right). \tag{5.8}$$

From Def. 5.2.2 and (5.7), it can be implied that if $\pi^* \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ is a SE, then the following holds, for all $k \in \mathcal{K}$,

$$\Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\left(\pi_k^*, \pi_{-k}^*\right) = 1. \tag{5.9}$$

Note that it can be stated that the set of equilibria of the game $\widehat{\mathcal{G}}$ is a subset of the set of equilibria of the mixed extension $\widehat{\mathcal{G}}'$, if we establish an injective map from the action set $\mathcal{A}_k$ to the set of vectors corresponding to the canonical basis of the space of N_k-dimensional vectors $\mathbb{R}^{N_k}$. For instance, let the n_k-th action of player k, i.e., $A_k^{(n_k)}$, be associated with the unitary vector $\mathbf{e}_{n_k}^{(N_k)} = \left(e_{n_k,1}^{(N_k)}, \ldots, e_{n_k,N_k}^{(N_k)}\right) \in \mathbb{R}^{(N_k)}$, where, all the components of the vector $\mathbf{e}_{n_k}^{(N_k)}$ are zero except its n_k-th component, which is unitary. The vector $\mathbf{e}_{n_k}^{(N_k)}$ represents a degenerated probability distribution, where the action $A_k^{(n_k)}$ is deterministically chosen. Using this argument, it becomes clear that every SE of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ is also a SE in the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$.

As we shall see in the next section, games in satisfaction form might not necessarily have a SE neither in pure strategies nor in mixed strategies. Thus, in the following we present a less restrictive notion of equilibrium to which we refer as epsilon-satisfaction equilibrium (ϵ-SE).

5.2.3 Epsilon-Satisfaction Equilibrium

At the SE of the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$, players choose their actions following a probability distribution such that only action profiles that allow all players to simultaneously satisfy their individual conditions are played with positive probability. This interpretation leads immediately to the conclusion that if there does not exist at least one action profile that allows all players to be simultaneously satisfied, then, there does not exist any SE in the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$. However, under certain conditions, it is always possible to build mixed strategies that allow players to be satisfied with a probability which is close to 1 , i.e., $1-\epsilon$, for a sufficiently small $\epsilon > 0$.

Definition 5.2.3 (Epsilon-Satisfaction Equilibrium) *Let $\epsilon \in]0,1]$. The mixed strategy profile $\pi^* \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$ is an epsilon-satisfaction equilibrium (ϵ-SE) of the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$, if for all $k \in \mathcal{K}$,*

$$\pi_k^* = \bar{\bar{f}}_k\left(\pi_{-k}\right), \tag{5.10}$$

where

$$\bar{\bar{f}}_k\left(\pi_{-k}^*\right) = \left\{\pi_k \in \triangle(\mathcal{A}_k) : \Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\left(\pi_k, \pi_{-k}^*\right) \geqslant 1-\epsilon\right\}. \tag{5.11}$$

From Def. 5.2.3, it can be implied that if the mixed strategy profile π^* is an ϵ-SE, it holds that,

$$\Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\left(\pi_k^*, \pi_{-k}^*\right) \geqslant 1-\epsilon. \tag{5.12}$$

That is, players are unsatisfied with probability ϵ. The relevance of the ϵ-SE is that it models the fact that players can be tolerant to a non-satisfaction level. At a given ϵ-SE, none of the players is interested in changing its mixed strategy profile as long as it is satisfied with a probability higher than certain threshold $1-\epsilon$. As we shall see, a game might not possess a SE neither in pure nor in mixed strategies, but it might possess an ϵ-SE.

The relevance of this result, for instance in terms of the example 5.2.1, is that when all the radio devices cannot be simultaneously satisfied, i.e., the required transmission rates cannot be simultaneously achieved, there may exist a way to let them achieve their required transmission rates with certain probability. In this case, ϵ corresponds to the maximum outage probability tolerated by all radio devices. Note that in none of the cases, we state that the ϵ is related to the minimum achievable outage probability. Here, radio devices are not interested in minimizing their outage probability, but rather to achieve a tolerable outage probability.

A thorough analysis on the existence and uniqueness of the SE in pure strategies and mixed strategies is presented in the next section. Similarly, the conditions for the existence of an ϵ-SE are also discussed.

5.3 Existence and Uniqueness of the Satisfaction Equilibrium

In this section, we study the existence and uniqueness of a satisfaction equilibrium in games in satisfaction form and their corresponding extension in mixed strategies. Particular attention is given to the existence of ϵ-SE in the case where there does not exist at least one SE neither in pure nor in mixed strategies.

5.3.1 Existence of SE in Pure Strategies

In order to study the existence of a SE in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$, let the correspondence $F : \mathcal{A} \rightarrow 2^{\mathcal{A}}$ be defined as follows:

$$F(\mathbf{a}) = f_1\left(\mathbf{a}_{-1}\right) \times \ldots \times f_K\left(\mathbf{a}_{-K}\right). \tag{5.13}$$

Then, a SE exists if and only if

$$\exists \mathbf{a} \in \mathcal{A} : \quad \mathbf{a} \in F(\mathbf{a}). \tag{5.14}$$

This formulation allows us to use existing fixed point (FP) theorems to provide sufficient conditions for the existence of the SE. For instance, in the case of compact and convex sets of actions, from Kakutani's FP theorem [40], we can write the following proposition.

Theorem 5.3.1 (Existence of the SE in compact games) *In the game* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$*, let the set of actions* $\mathcal{A}$ *be a non-empty, convex and compact set. Let also the correspondence* F *have a closed graph and for all* $\mathbf{a} \in \mathcal{A}$*, let* $F(\mathbf{a})$ *be non-empty and convex. Then, the game* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *has at least one SE.*

In the case of finite sets of actions, one can rely on the fixed point theorem of Knaster and Tarski [42] to state the following theorem.

Theorem 5.3.2 (Existence of SE in finite games) *Consider the game* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *and let the set* $\mathcal{A}$ *have a binary relation denoted by* $\preceq$*. Let also*
(i) $V = \langle \mathcal{A}, \preceq \rangle$ *be a complete lattice;*
(ii) $F(\mathbf{a})$ *be non-empty for all* $\mathbf{a} \in \mathcal{A}$*;*
(iii) the correspondence F *in* (5.13) *satisfies that* $\forall (\mathbf{a}, \mathbf{a}') \in \mathcal{A}$*, such that* $\mathbf{a} \preceq \mathbf{a}'$*, it holds that*

$$\forall (\mathbf{b}, \mathbf{b}') \in F(\mathbf{a}) \times F(\mathbf{a}'), \quad \mathbf{b} \preceq \mathbf{b}'. \tag{5.15}$$

Then the game has at least one SE in pure strategies.

Note that both theorem 5.3.1 and theorem 5.3.2 require the following assumptions that for all $\mathbf{a} \in \mathcal{A}$, the set $F(\mathbf{a})$ is non-empty, i.e.,

$$\forall k \in \mathcal{K} \text{ and } \forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}, \exists a_k \in \mathcal{A}_k : \, a_k \in f_k(\mathbf{a}_{-k}). \tag{5.16}$$

However, this condition is highly demanding. In practical terms, it implies that for all the radio devices, there always exists a transmit/receive configuration such that their QoS requirement are satisfied. Nonetheless, this is not always the case.
In the following, we study the existence of an equilibrium in mixed strategies.

5.3.2 Existence of the SE in Mixed Strategies

As in the case of pure strategies, the condition for the existence of a SE in the mixed extension $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$ boils down to the study of a fixed point inclusion. Let the correspondence $\bar{F} : \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K) \rightarrow 2^{\triangle(\mathcal{A}_1)\times\ldots\times\triangle(\mathcal{A}_K)}$ be defined as follows:

$$\bar{F}(\pi) = \bar{f}_1(\pi_{-1}) \times \ldots \times \bar{f}_K(\pi_{-K}). \tag{5.17}$$

Then, a SE exists if and only if

$$\exists \pi \in \triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K) : \quad \pi \in \bar{F}(\pi). \tag{5.18}$$

Thus, all the results of fixed point theory [11], in the case of the compact and convex sets, are valid for the study of the existence of the SE in the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \{\triangle(\mathcal{A}_k)\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$. Nonetheless, some results are immediate from Def. 5.2.2. For instance, note that if a game in satisfaction form does not have a SE in pure strategies, then, it does not have a SE in mixed strategies neither. This is basically due to the fact that players mix only the actions that guarantee their satisfaction with probability one. That is, player k mixes a subset of its actions $\mathcal{A}'_k \subseteq \mathcal{A}_k$, i.e., $\forall a_k \in \mathcal{A}'_k$, $\pi_{k,a_k} > 0$, only if the following condition holds,

$$\begin{aligned} \forall a_k^* \in \mathcal{A}'_k, \qquad \Pr\left(a_k^* \in f_k\left(\mathbf{a}_{-k}\right)\right)\pi_{-k} &= 1 \qquad (5.19)\\ \sum_{\forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}} \mathbb{1}_{\left\{a_k^* \in f_k(\mathbf{a}_{-k})\right\}} \prod_{j \in \mathcal{K}\setminus\{k\}} \pi_{j,a_j} &= 1. \qquad (5.20) \end{aligned}$$

This implies that player k assigns a strictly positive probability to more than one action, i.e., it plays strictly mixed strategies, only if such a set of actions guarantees its satisfaction for all the action profiles $\mathbf{a}_{-k} \in \mathcal{A}_{-k}$, which are played with non-zero probability. This reasoning might imply that, there might exist several SE in pure strategies but no SE in strictly mixed strategies in the game $\widehat{\mathcal{G}}' = \left\{\mathcal{K}, \left\{\triangle\left(\mathcal{A}_k\right)\right\}_{k\in\mathcal{K}}, \left\{\bar{f}_k\right\}_{k\in\mathcal{K}}\right\}$.

5.3.3 Existence of the Epsilon-Satisfaction Equilibrium

As shown in the previous subsection, the existence of at least one SE in the extension in mixed strategies of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$ remains very strict. Indeed, the game has a SE in mixed strategies if and only if it has a SE in pure strategies. On the contrary, the existence of at least one ϵ-SE is less strict and it does not require the existence of a pure SE. A sufficient and necessary condition for the existence of at least one ϵ-SE is the following.

Proposition 5.3.1 *Let $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$ be a finite game in satisfaction form. Then, if the following condition holds,*

$$\forall k \in \mathcal{K}, \exists \mathbf{a} \in \mathcal{A}: \quad a_k \in f_k\left(\mathbf{a}_{-k}\right), \qquad (5.21)$$

there always exists a strategy profile $\pi^* \in \triangle\left(\mathcal{A}_1\right) \times \ldots \times \triangle\left(\mathcal{A}_K\right)$ and an $1 > \epsilon > 0$, such that, π^* is an ϵ-SE.

Proof: *Assume that the condition (5.21) holds. Then, for all $j \in \mathcal{K}$, it holds that the set*

$$\mathcal{A}_j^* = \left\{\mathbf{a} \in \mathcal{A} : a_j \in f_j\left(\mathbf{a}_{-j}\right)\right\} \qquad (5.22)$$

is non-empty. Denote by $\mathbf{a}_j^* = \left(a_{j,1}^*, \ldots, a_{j,K}^*\right)$ a particular element of the set $\mathcal{A}_j^*$. Any mixed strategy $\pi^+ \in \triangle\left(\mathcal{A}_1\right) \times \ldots \times \triangle\left(\mathcal{A}_K\right)$, such that

$$\forall (j,k) \in \mathcal{K}^2, \quad \pi^+_{k,a_{j,k}^*} > 0 \qquad (5.23)$$

guarantees that $\forall j \in \mathcal{K}$, the action $\mathbf{a}_j^*$ is played with non-zero probability, thus,

$$\forall k \in \mathcal{K}, \quad \Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\pi^+ = \sum_{\mathbf{a}\in\mathcal{A}} \mathbb{1}_{\left\{a_k \in f_k(\mathbf{a}_{-k})\right\}} \prod_{j=1}^{K} \pi^+_{j,a_j} = \epsilon_k, \qquad (5.24)$$

where $1 \geqslant \epsilon_k \geqslant \prod_{j=1}^{K} \pi^+_{j,a_{k,j}^*} > 0$, which proves the existence of a mixed strategy profile such that,

$$\forall k \in \mathcal{K}, \quad \Pr\left(a_k \in f_k\left(\mathbf{a}_{-k}\right)\right)\pi^+ \geqslant 1 - \epsilon, \qquad (5.25)$$

where, $\epsilon = 1 - \max_{j \in \mathcal{K}} \epsilon_j$*, which completes the proof.* □

Note that for all $k \in \mathcal{K}$, the condition (5.21) only requires the existence of at least one action profile where player k is satisfied, which is less restrictive than conditions (5.14) and (5.18). Note also that, as long as (5.21) holds, a simple uniform distribution over each individual set of actions $\mathcal{A}_k$ is an ϵ-SE, where $\epsilon = 1 - \prod_{j=1}^{K} \frac{1}{N_j}$.

5.3.4 Uniqueness of the SE

In general, it is difficult to provide the conditions to observe a unique SE for a general set of correspondences $\{f_k\}_{k \in \mathcal{K}}$. As we shall see in Sec. 5.7, the set of SE is often non-unique in games modeling decentralized self-configuring wireless networks, and thus, an equilibrium selection process might be required. In Sec. 5.5, we propose a methodology for equilibrium selection, but first, we focus on establishing the differences between the notion of SE and the notions of NE and GNE.

5.4 Satisfaction Equilibrium and other Equilibrium Concepts

In the following, we highlight the main differences between the SE and other equilibrium notions such as NE and GNE. However, before we start, we point out the differences between the normal form and the satisfaction form formulations.

5.4.1 Games in Normal Form and Satisfaction Form

The main difference between the normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}}\right\}$ and satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}}\right\}$ is that the former defines a utility function u_k in the sense of Neumann - Morgenstern [68], i.e., given an action profile $\mathbf{a}_{-k} \in \mathcal{A}_{-k}$, player k can rank any pair of its actions $(a_k, a'_k) \in \mathcal{A}_k^2$ such that either $u_k(a_k, \mathbf{a}_{-k}) < u_k(a'_k, \mathbf{a}_{-k})$, $u_k(a_k, \mathbf{a}_{-k}) = u_k(a'_k, \mathbf{a}_{-k})$ or $u_k(a_k, \mathbf{a}_{-k}) > u_k(a'_k, \mathbf{a}_{-k})$. In the latter, player k determines only whether an action satisfies its individual conditions or not, i.e., $a_k \in f_k(\mathbf{a}_{-k})$ or $a_k \notin f_k(\mathbf{a}_{-k})$, respectively.

In the following, we present a simple example that allows us to identify the differences between modeling the QoS problem using a normal-form formulation and a satisfaction formulation. Here, we model the scenario described in Ex. 5.2.1 by a game in satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{f_k\}_{k \in \mathcal{K}}\right\}$, where f_k is defined by (5.4), and a game in normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{v_k\}_{k \in \mathcal{K}}\right\}$. In the normal form, we assume that a player gets 1 if it is able to achieve satisfaction or 0 otherwise. Hence, the function $v_k : \mathcal{A}_1 \times \ldots \times \mathcal{A}_K \to \{0, 1\}$ is defined as follows, for all $k \in \mathcal{K}$,

$$v_k(a_k, \mathbf{a}_{-k}) = \mathbb{1}_{\{a_k \in f_k(\mathbf{a}_{-k})\}}. \tag{5.26}$$

Now, we compare both the set of SE $\mathcal{A}_{\mathrm{SE}}$ of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ and the set of NE $\mathcal{A}_{\mathrm{NE}}$ of the game $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$. Note that from Def. 5.4.1 and Def. 5.2.1, it can be immediately implied that any SE of the game $\widehat{\mathcal{G}}$ is an NE of the game $\mathcal{G}$. This is basically, because at the SE, all players obtain a unitary utility, and since the range of the utility function is binary $\{0,1\}$, no other action is able to give a higher utility. The converse is not true, that is, an NE of the game $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$ is not necessarily a SE of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$. Consider for instance the game realization ($K = 2$, $N_1 = N_2 = 2$) in Fig. 5.1. Note that therein, the game $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$ has 2 NE in pure strategies, which are the action profiles $(A_1^{(2)}, A_2^{(1)})$ and $(A_1^{(1)}, A_2^{(2)})$, while the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ has only one SE, which is the action profile $(A_1^{(1)}, A_2^{(2)})$. This simple example shows that, the game formulation following

$P_1 \backslash P_2$	$A_2^{(1)}$	$A_2^{(2)}$
$A_1^{(1)}$	$(0,0)$	$(1,1)$
$A_1^{(2)}$	$(1,0)$	$(0,0)$

Figure 5.1: Game in normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$, with $\mathcal{K} = \{1,2\}$, $\mathcal{A}_k = \{A_k^{(1)}, A_k^{(2)}\}$, for all $k \in \mathcal{K}$. Player 1 chooses rows and player 2 chooses columns. In a pair $(v_1, v_2) \in \{0,1\}^2$, v_1 and v_2 are the utilities obtained by player 1 and 2, respectively.

the idea of a utility function of the form $v_k : \mathcal{A} \rightarrow \mathbb{R}_+$, for all $k \in \mathcal{K}$, might lead to equilibria where not all the players are satisfied. This shows that games in normal-form do not properly model the case where players are interested only in the satisfaction of individual conditions. We conclude the comparison between the games $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{v_k\}_{k\in\mathcal{K}}\right\}$ and $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$, by establishing the following condition between their sets of equilibria.

$$\mathcal{A}_{\mathrm{SE}} \subseteq \mathcal{A}_{\mathrm{NE}} \subseteq \mathcal{A}. \tag{5.27}$$

This confirms the intuition that the notion of SE is more restrictive than the notion of NE, that is, an SE in the game $\widehat{\mathcal{G}}$ is an NE in the game $\mathcal{G}$, where all players are satisfied.

5.4.2 Satisfaction Equilibrium and Nash Equilibrium

The NE in pure strategies in the context of games in normal form [31] can be defined as follows.

Definition 5.4.1 (Nash Equilibrium in PS [66]) *Consider a game in normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right\}$. An action profile $\boldsymbol{a} \in \mathcal{A}$ is an NE in pure strategies if it satisfies, for all $k \in \mathcal{K}$ and for all $\mathbf{a}'_k \in \mathcal{A}_k$,*

$$u_k(a_k, \mathbf{a}_{-k}) \geqslant u_k(a'_k, \mathbf{a}_{-k}). \tag{5.28}$$

Note that the definition of NE (Def. 5.4.1) can be obtained from the definition of SE (Def. 5.2.1) by assuming that, for all $k \in \mathcal{K}$, the satisfaction correspondence f_k is defined as follows,

$$f_k\left(\mathbf{a}_{-k}\right) = \arg \max_{a_k^* \in \mathcal{A}_k} u_k\left(a_k^*, \mathbf{a}_{-k}\right). \tag{5.29}$$

The satisfaction correspondence f_k as defined in (5.29) is known in the game theoretic literature as the best response correspondence [31]. Then, under this formulation, the set of SE of the game in satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ is identical to the set of NE of the game in normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right\}$. This reasoning might lead us to think that the satisfaction form as well as the notion of SE are generalizations of the classical normal form and the notion of Nash equilibrium [66], respectively.

5.4.3 Satisfaction Equilibrium and Generalized Nash Equilibrium

The GNE in pure strategies (PS) in games in normal form with constrained set of actions, as introduced by Debreu in [24] and later by Rosen in [92], can be defined as follows.

Definition 5.4.2 (Generalized NE in PS [24]) *An action profile $\boldsymbol{a}^* \in \mathcal{A}$ is a generalized Nash equilibrium (GNE) of the game $\hat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ if and only if*

$$\forall k \in \mathcal{K}, \quad a_k^* \in f_k\left(\mathbf{a}_{-k}^*\right) \text{ and}$$

$$\forall a_k \in f_k\left(\mathbf{a}_{-k}^*\right),\ u_k(a_k^*, \mathbf{a}_{-k}^*) \geqslant u_k(a_k, \mathbf{a}_{-k}^*).$$

Note that the definition of SE (Def. 5.2.1) can be obtained from the definition of GNE (Def. 5.4.2) by assuming the following condition, $\forall k \in \mathcal{K}$ and $\forall \mathbf{a} \in \mathcal{A}$

$$u_k(a_k, \mathbf{a}_{-k}) = c, \quad \text{with } c \in \mathbb{R}_+. \tag{5.30}$$

Under assumption (5.30), the set of GNE of the game in normal form with constrained set of actions $\hat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ and the set of SE of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ in satisfaction form are identical. This observation does not necessarily imply that the satisfaction form as well as the notion of SE are particular cases of the classical normal form with constrained set of actions and the notion of GNE [24], respectively. Note that, in the game $\hat{\mathcal{G}}$ the action profile $\mathbf{a} \in \mathcal{A}$ can be a possible game outcome, if and only if, it satisfies that for all $k \in \mathcal{K}$, $a_k \in f_k\left(\mathbf{a}_{-k}\right)$. Conversely, in the game $\widehat{\mathcal{G}}$ any action profile of the set $\mathcal{A}$ is a possible game outcome. That is, the game formulation $\widehat{\mathcal{G}}$ is not a formulation with constrained set of actions.
In the following, we compare the set of equilibria of both games $\hat{\mathcal{G}}$ and $\widehat{\mathcal{G}}$, for a general definition of the utility functions u_k, for all $k \in \mathcal{K}$. Let the sets of GNE of the game $\hat{\mathcal{G}}$ and the set of SE of the game $\widehat{\mathcal{G}}$ be denoted by $\mathcal{A}_{\text{GNE}}$ and $\mathcal{A}_{\text{SE}}$,

respectively. Now, note that from Def. 5.4.2 and Def. 5.2.1, it follows that any GNE in $\hat{\mathcal{G}}$ is a SE in $\widehat{\mathcal{G}}$, i.e.,

$$\mathcal{A}_{\mathrm{GNE}} \subseteq \mathcal{A}_{\mathrm{SE}} \subseteq \mathcal{A}. \tag{5.31}$$

The condition in (5.31) verifies the intuition that the notion of SE in games in satisfaction form, is less restrictive than the notion of GNE in games in normal form with constrained action sets. Note also that from Def. 5.2.1, it might be implied that several SE might exist, while no GNE necessarily exists. This is basically due to the fact that the existence of a GNE depends on both the functions u_k and f_k, while the existence of a SE depends uniquely on the correspondences f_k, with $k \in \mathcal{K}$. Conversely, the existence of a GNE implies the existence of a SE.

5.5 Equilibrium Selection and Efficient Satisfaction Equilibria

In this section, we tackle the equilibrium selection process when a game in satisfaction form possesses several (satisfaction) equilibria. We start our analysis pointing the fact that at the SE, all radio devices are able to provide the required QoS. Hence, none of them has an interest in unilaterally changing the actual transmit configuration. However, using a higher transmit power level or using a more complex modulation scheme (e.g., in the sense of the size of the constellation) might require a higher energy consumption and thus, reduce the battery life time of the transmitters. In this scenario, one might imply that radio devices are interested in satisfying their required QoS with the lowest effort. Here, we can express the effort, for instance, in terms of energy consumption or signal processing complexity. Any action profile which allows all the players to be satisfied with the lowest effort is an efficient satisfaction equilibrium.

In the sequel of this section, we formulate a game where its set of (generalized Nash) equilibria coincide with the notion of efficient satisfaction equilibrium (ESE). Later, we analyze the existence and uniqueness of the equilibrium of such game.

5.5.1 Games in Efficient Satisfaction Form

In the following, we assume that ever player k is able to built a binary relation denoted by $\prec_k$ on its own set of actions $\mathcal{A}_k$ such that the partially ordered set $\langle \mathcal{A}_k, \prec_k \rangle$ is a complete lattice. The ordering is as follows: $\forall (a_k, a_k') \in \mathcal{A}^2$, $a_k \prec_k a_k'$ implies that a_k requires less effort than a_k' when it is played by player k. An important remark here is that, the effort assigning process, i.e., the definition of the binary relation $\prec_k$, is independent of the other players' choice.

The game where each player aims to satisfy its QoS with the minimum effort can be formulated as a game in normal form with constrained set of actions,

$$\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}} \right). \tag{5.32}$$

Here, for all $k \in \mathcal{K}$, the utility function $c_k : \mathcal{A}_k \to [0,1]$ of player k satisfies the following condition, $\forall (a_k, a'_k) \in \mathcal{A}_k^2$,

$$a_k \prec_k a'_k \Longleftrightarrow c_k\left(a_k\right) < c_k\left(a'_k\right). \tag{5.33}$$

We refer to this particular game formulation and its set of (generalized) NE as the efficient satisfaction form (ESF) and set of ESE, respectively.

Definition 5.5.1 (Efficient SE) *Consider the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right\}$ in satisfaction form. An action profile $\mathbf{a}^* \in \mathcal{A}$ is an efficient SE, with respect to the cost functions $\{c_k\}_{\forall k\in\mathcal{K}}$, if it is a (generalized) NE of the game in normal form with constrained set of actions $\tilde{\mathcal{G}} = \left(\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{c_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right)$, i.e.,*

$$\forall k \in \mathcal{K}, \quad a_k^* \in f_k\left(\mathbf{a}_{-k}^*\right) \ \text{and}$$

$$\forall a_k \in f_k\left(\mathbf{a}_{-k}^*\right), \ c_k(a_k^*) \leqslant c_k(a_k).$$

It is important to note that in the game $\tilde{\mathcal{G}} = \left(\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{c_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right)$, the competitive interaction between all players is not modeled by the cost functions $\{c_k\}_{k\in\mathcal{K}}$. For instance, the cost function of player k, c_k, depends only on its chosen action a_k. In this game formulation $\tilde{\mathcal{G}}$, the interaction between players is modeled by the correspondence f_k, which is defined over the set of action profiles $\mathcal{A}_{-k}$.
An important remark on Def. 5.5.1 is that if all players assign the same cost (or effort) to all their actions, then the sets of ESE and the set of SE are identical. This implies that the interest of the formulation $\tilde{\mathcal{G}}$ is precisely that players can differentiate the effort of playing one action or another in order to select one (satisfaction) equilibrium among all the existing equilibria of the game $\widehat{\mathcal{G}}$. Thus, the existence and uniqueness of this efficient SE plays an important role in the equilibrium selection. We analyze this two properties in the sequel of this section.

5.5.2 Existence of an ESE

Before giving a formal result on the existence of the ESE, we introduce a generalization of a class of games known as exact potential games (PG) [64]. We refer to this new class of games as constrained exact potential games. First, consider a game in normal form with constrained strategies and denote it by $\tilde{\mathcal{G}} = \left(\mathcal{K}, \left\{\mathcal{A}_k\right\}_{k\in\mathcal{K}}, \left\{c_k\right\}_{k\in\mathcal{K}}, \left\{f_k\right\}_{k\in\mathcal{K}}\right)$. Let the set $\mathcal{F}_k \subset \mathcal{A}$ be the graph of the correspondence f_k, hence,

$$\mathcal{F}_k = \left\{(a_k, \mathbf{a}_{-k}) \in \mathcal{A} : a_k \in f_k\left(\mathbf{a}_{-k}\right)\right\}. \tag{5.34}$$

The set $\mathcal{F}_k$ determines the action profiles which can be observed as outcomes of the game $\tilde{\mathcal{G}}$, when only player k is allowed to play given any action profile $\mathbf{a}_{-k}$ for which the set $f_k\left(\mathbf{a}_{-k}\right)$ is not empty. Following this reasoning, the set of all possible outcomes of the game $\widehat{\mathcal{G}}$ corresponds to the following set

$$\mathcal{F} = \bigcap_{j=1}^{K} \mathcal{F}_j, \tag{5.35}$$

which is the set of action profiles such that $\forall \mathbf{a} \in \mathcal{F}$, it holds that $\forall k \in \mathcal{K}$, $a_k \in f_k(\mathbf{a}_{-k})$. However, unilateral deviations of a set of players from any action profile $\mathbf{a} \in \mathcal{F}$ might lead to action profiles which do not belong to $\mathcal{F}$. The following set

$$\widehat{\mathcal{F}} = \bigcup_{j=1}^{K} \mathcal{F}_j, \tag{5.36}$$

contains all possible unilateral deviations one can observe from any action in the set $\mathcal{F}$. Using both sets $\mathcal{F}$ and $\widehat{\mathcal{F}}$, we introduce the definition of exact constrained potential game.

Definition 5.5.2 (Exact Constrained PG (ECPG)) *Any game in normal form with constrained set of actions $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ is an exact constrained potential game (ECPG) if there exists a function $\phi : \widehat{\mathcal{F}} \to \mathbb{R}$ such that for all $\mathbf{a} \in \widehat{\mathcal{F}}$, it holds that, for all $k \in \mathcal{K}$ and for all $a'_k \in f_k(\mathbf{a}_{-k})$,*

$$u_k(a_k, \mathbf{a}_{-k}) - u_k(a'_k, \mathbf{a}_{-k}) = \phi(a_k, \mathbf{a}_{-k}) - \phi(a'_k, \mathbf{a}_{-k}).$$

Before we continue, we clearly state that not all the the properties of potential games [64] hold for the constrained potential games. For instance, not all exact constrained PG have an equilibrium. In the following, we introduce two results regarding the existence of an equilibrium in pure strategies in ECPG.

Theorem 5.5.3 (Existence of an equilibrium in ECPG) *The finite exact constrained potential game $\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$, with potential function $\phi : \widehat{\mathcal{F}} \to \mathbb{R}_+$, has at least one equilibrium in pure strategies, if the sets $\mathcal{F}$ (5.35) and $\widehat{\mathcal{F}}$ (5.36) are non-empty and identical.*

Proof: *By assumption, the set $\mathcal{F}$ (5.35) is non-empty. Thus, there exists at least one feasible outcome $\mathbf{a}^* \in \mathcal{F}$ for the game $\mathcal{G}$. Now, for all $k \in \mathcal{K}$, any unilateral deviation of player k from an action profile $\mathbf{a}^*$ leads to an action profile of the form $\left(a_k, \mathbf{a}^*_{-k}\right) \in \widehat{\mathcal{F}}$. Similarly, by assumption, both sets $\mathcal{F}$ and $\widehat{\mathcal{F}}$ are identical, thus, any unilateral deviation from a feasible action profile is also a feasible action profile. Now, without any loss of generality, let the elements of the sets $\mathcal{F} = \left\{\tilde{A}^{(1)}, \ldots, \tilde{A}^{(N)}\right\}$ be indexed following any particular order such that the following holds,*

$$\phi\left(\tilde{A}^{(1)}\right) \leqslant \phi\left(\tilde{A}^{(2)}\right) \leqslant \ldots \leqslant \phi\left(\tilde{A}^{(N)}\right), \tag{5.37}$$

with $N = |\mathcal{F}|$. Thus, from Def. 5.4.2, it holds that $\tilde{A}^{(N)}$ is an equilibrium of the game $\mathcal{G}$, which completes the proof. □

Theorem 5.5.4 (Existence of Equilibrium in ECPG) *Let*

$$\mathcal{G} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$$

be an exact constrained potential game, with potential function $\phi : \widehat{\mathcal{F}} \to \mathbb{R}_+$ continuous over a finite dimensional linear space containing the set $\widehat{\mathcal{F}}$ (5.35). Let the sets $\mathcal{F}$ and $\widehat{\mathcal{F}}$ (5.36) be two identical non-empty compact and convex sets. Then, the game $\mathcal{G}$ has at least one equilibrium in pure strategies.

Proof: *By assumption, the set* $\mathcal{F}$ (5.35) *is non-empty. Thus, there exists at least one feasible outcome* $\mathbf{a}^* \in \mathcal{F}$ *for the game* $\mathcal{G}$*. Now, for all* $k \in \mathcal{K}$*, any unilateral deviation of player* k *from an action profile* $\mathbf{a}^*$ *leads to an action profile of the form* $\left(a_k, \mathbf{a}^*_{-k}\right) \in \widehat{\mathcal{F}}$*. Similarly, by assumption, both sets* $\mathcal{F}$ *and* $\widehat{\mathcal{F}}$ *are identical convex and compact, thus, any unilateral deviation from a feasible action profile is also a feasible action profile. Now, from the fact that any continuous function defined over a compact and convex set achieves a maximum and any maximum of the potential function is an equilibrium, it follows that the game* $\mathcal{G}$ *has always at least one equilibrium, which completes the proof.* □

Now, using Def. 5.5.2, we introduce the following proposition.

Proposition 5.5.1 *Every game* $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ *in efficient-satisfaction form is an exact constrained potential game, with potential function* $\phi : \mathcal{A} \rightarrow \mathbb{R}$*, such that,* $\forall \mathbf{a} \in \mathcal{A}$*,*

$$\phi\left(\mathbf{a}\right) = \sum_{k=1}^{K} c_k\left(a_k\right). \tag{5.38}$$

Note that Prop. 5.5.1 is an immediate result from Def. 5.5.2. The following two corollaries are immediately obtained from both Th. 5.5.3 and Th. 5.5.4, respectively.

Corollary 5.5.5 (Existence of the ESE) *A finite game in efficient satisfaction form* $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$*, with non-empty and identical sets* $\mathcal{F}$ (5.35) *and* $\widehat{\mathcal{F}}$ (5.36) *, always has at least one (efficient satisfaction) equilibrium.*

Corollary 5.5.6 (Existence of the ESE) *The game in efficient satisfaction form* $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$*, with the sets* $\mathcal{F}$ (5.35) *and* $\widehat{\mathcal{F}}$ (5.36) *identical, non-empty, compact and convex, and* c_k *continuous over a finite dimensional linear space containing the set* $\mathcal{F}$*, for all* $k \in \mathcal{K}$*, always has at least one (efficient satisfaction) equilibrium.*

In corollary 5.5.5 and corollary 5.5.6, we have established sufficient but not necessary conditions for the existence of an efficient SE. In the following, we study the uniqueness of such ESE, when it exists.

5.5.3 Uniqueness of the ESE

In the following, we study the uniqueness of the equilibrium of the potential game with constrained strategies $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$. For doing so, we analyze the auxiliary game defined by the game in normal form with constrained action sets $\tilde{\mathcal{G}}' = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{\phi\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$. The interest on adopting this approach stems from the fact that the set of equilibria of both games are identical (Prop. 5.5.1) and the existence of the potential ϕ facilitates our analysis. Following this reasoning, a strategy profile $\mathbf{a}^* \in \mathcal{A}$ is an equilibrium of the game $\tilde{\mathcal{G}}' = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{\phi\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$, and thus, an ESE of $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$, if

$$\forall k \in \mathcal{K}, \quad \mathbf{a}^*_k \in \arg \max_{a_k \in f_k(\mathbf{a}_{-k})} \phi\left(a_k, \mathbf{a}^*_{-k}\right). \tag{5.39}$$

Hence, the proof of the uniqueness of the equilibria of the game $\tilde{\mathcal{G}}$, in the case of compact and convex action sets, reduces to prove that the optimization problem (5.39) has a unique solution. Hence, we state the following proposition.

Proposition 5.5.2 (ESE in compact set of actions) *The game in efficient satisfaction form* $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$, *with the sets* $\mathcal{F}$ (5.35) *and* $\widehat{\mathcal{F}}$ (5.36) *identical, non-empty, compact and convex and the functions* $\{c_k\}_{k\in\mathcal{K}}$ *continuous and strictly convex over a finite dimensional linear space containing the set* $\mathcal{F}$ *has a unique efficient satisfaction equilibrium in pure strategies.*

Proof: *For all action profile* $\mathbf{a}^* \in \mathcal{F}$ *and for all* $k \in \mathcal{K}$ *any unilateral deviation of player* k *from an action profile* $\mathbf{a}^*$ *leads to the action profile of the form* $\left(a_k, \mathbf{a}^*_{-k}\right)$, *with* $a_k \in f_k\left(\mathbf{a}_{-k}\right)$. *Now, since both sets* $\mathcal{F}$ *and* $\widehat{\mathcal{F}}$ *are identical convex and compact, it holds that* $\left(a_k, \mathbf{a}^*_{-k}\right) \in \mathcal{F}$. *This implies that any unilateral deviation from a feasible action profile yields a feasible action profile. Now, from the fact that the potential is continuous and strictly convex (sum of continuos and strictly convex functions* $\{c_k\}_{\forall k\in\mathcal{K}}$,*) over the set* $\mathcal{F}$, *it holds that there exists a minimum of the potential and it is unique, which completes the proof.* □

In the case of discrete sets of actions, it is easy to show that the ESE might not be unique. In the following, we use some tools from graph theory to determine the number of ESE which a given game in efficient-satisfaction form $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ can possess. We start by indexing the elements of the action set $\mathcal{A}$ in any given order using the index $n \in \mathcal{I} = \{1, \ldots, |\mathcal{A}|\}$. Denote by $\mathbf{a}^{(n)} = \left(a_1^{(n)}, \ldots, a_K^{(n)}\right)$ the n-th element of the action set $\mathcal{A}$. Consider that each action profile $\mathbf{a}^{(n)}$ is associated with a vertex x_n in a given directed graph G. There exists an arc from vertex x_n to another vertex x_m, if the action profile represented by the latter $\mathbf{a}^{(m)}$ can be obtained from the former $\mathbf{a}^{(n)}$ by changing the action of only one player and lower potential (sum of efforts) is obtained. For instance, if the unique deviator is player k, then, $a_k^{(m)} \in f_k\left(\mathbf{a}_{-k}^{(n)}\right)$ and $\phi\left(\mathbf{a}^{(n)}\right) > \phi\left(\mathbf{a}^{(m)}\right)$. More precisely, the graph G can be defined by the pair $G = (\mathcal{X}, \mathbf{B})$, where the set $\mathcal{X} = \left\{x_1, \ldots, x_{|\mathcal{A}|}\right\}$ (nodes) contains the nodes representing the action profiles in the set $\mathcal{A}$ and $\mathbf{B}$ (edges) is a non-symmetric matrix with dimensions $|\mathcal{A}| \times |\mathcal{A}|$ and entries defined as follows $\forall (n,m) \in \mathcal{I}^2$ and $n \neq m$,

$$b_{n,m} = \begin{cases} 1 & \text{if} \quad (i)\ \exists!\, k \in \mathcal{K} :\ a_k^{(n)} \neq a_k^{(m)},\ \text{and}\ a_k^{(m)} \in f_k\left(a_{-k}^{(n)}\right) \\ & \phantom{\text{if}} \quad (ii)\ \phi\left(\mathbf{a}^{(m)}\right) < \phi\left(\mathbf{a}^{(n)}\right) \\ 0 & \text{otherwise}\ , \end{cases} \tag{5.40}$$

and $b_{i,i} = 0$ for all $i \in \mathcal{I}$.

A realistic assumption is to consider that for any pair of action profiles $\mathbf{a}^{(n)}$ and $\mathbf{a}^{(m)}$ which are adjacent, we have that $\phi(\mathbf{a}^{(n)}) \neq \phi(\mathbf{a}^{(m)})$. This is because players assign different effort values to their actions. From the definition of the matrix $\mathbf{B}$, we have that a necessary and sufficient condition for a vertex x_n to represent an ESE action profile is to have a null out-degree in the oriented graph G, i.e., there are no outgoing edges from the node x_n (sink vertex). Finally, one can conclude that

determining the set of ESE in the game $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right)$ boils down to identifying all the sink vertices in the oriented graph G. That is, exploiting the fact that, if the n-th row of the matrix $\mathbf{B}$ is a null-vector, then the action $\mathbf{a}^{(n)}$ is an ESE of the game $\tilde{\mathcal{G}}$. Interestingly, a particular case arises when the resulting graph is an edgeless graph, i.e., the corresponding matrix $\mathbf{B}$ is a null matrix. In this case, the set of SE would be identical to the set of ESE, which implies that the idea of associating an effort to each action is not enough to select an ESE among the set of SE. In any case, determining the exact set of SE would require the analysis of the matrix $\mathbf{B}$, which might be highly demanding and requires complete information.
In the following, we focus on designing behavioral rules for the radio devices in order to let them to learn one satisfaction equilibrium in decentralized self-configuring networks.

5.6 Learning Satisfaction Equilibrium

In this section, we study a behavioral rule that allows radio devices to learn a satisfaction equilibrium in a fully decentralized fashion. Here, the underlying assumption is that players do not need to observe the value of its achieved utility, i.e., transmission rate, energy efficiency, etc., but only to know whether they are satisfied or not at each stage of the learning process, which implies a 1-bit length message exchange between the corresponding transmitter-receiver pairs. In the following, we formulate the corresponding learning problem and later, we introduce the behavioral rules that allow players to learn the SE.

5.6.1 The Learning Problem Formulation

We describe the SE learning process in terms of elements of the game in satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ as follows. Assume that time is divided in intervals and denote each interval by the index $n \in \mathbb{N}$. Each interval ends when each player has played at most once. Denote the action taken by player k at interval n by $a_k(n)$. At each interval n, player k observes whether it is satisfied or not, i.e., it observes a binary variable

$$\tilde{v}_k(n) = \mathbb{1}_{\{a_k(n)\in f_k(\mathbf{a}_{-k}(n))\}}. \tag{5.41}$$

Note that this observation requires only a 1-bit message exchange between the corresponding transmitter and receiver pair. Our intention is to learn at least one SE by letting the players to interact following particular behavioral rules. We say that players learn an equilibrium in pure strategies if, after a given finite number of time intervals, all players have chosen an action which achieves satisfaction, and thus, no other action update takes place. We say that players learn an ϵ-SE if during a large observation period of T intervals, player k has been satisfied during at least t intervals, with $\frac{t}{T} > \epsilon$, for all $k \in \mathcal{K}$.

5.6.2 Learning the SE in Pure Strategies

Before we present the behavioral rule which allows players to achieve one of the equilibrium of the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$, we state the following hypothesis:

(i) The game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ has at least one SE in pure strategies.

(ii) For all $k \in \mathcal{K}$, it holds that $\forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}$, there exists at least one $a_k \in \mathcal{A}_k$ such that $a_k \in f_k(\mathbf{a}_{-k})$.

(iii) The sets $\mathcal{K}$ and $\{\mathcal{A}_k\}_{k\in\mathcal{K}}$, are finite.

The first hypothesis ensures that the SE learning problem is well-posed, i.e., radio devices are assigned a feasible task. The second hypothesis refers to the fact that, each radio device is always able to find a transmit/receive configuration such that it can satisfy its individual QoS requirement given the transmit/receive configuration of all the other radio devices. The third hypothesis is considered in order to ensure that our algorithm is able to converge in finite time.
Under the assumption that all hypothesis hold, each player chooses its own action as follows. The first action of player k, denoted by $a_k(0)$, is taken following an arbitrary probability distribution $\hat{\pi}_k(0) \in \triangle(\mathcal{A}_k)$. Often, such a probability $\hat{\pi}_k(0)$ is the uniform probability distribution. At time interval $n > 0$, player k changes its action if and only if it is not satisfied, i.e, $\tilde{v}_k(n-1) = 0$. In this case, the next action is chosen following a probability distribution $\hat{\pi}_k(n)$. If player k is satisfied i.e, $\tilde{v}_k(n-1) = 1$, then, it keeps playing the same action. Hence, we can write that,

$$a_k(n) = \begin{cases} a_k(n-1) & \text{if} \quad \tilde{v}_k(n-1) = 1 \\ a_k(n) \sim \hat{\pi}_k(n) & \text{if} \quad \tilde{v}_k(n-1) = 0 \end{cases}. \tag{5.42}$$

The behavioral rule (5.42) was first proposed in [94]. Therein, two particular ways for building the probability distribution $\hat{\pi}_k(n)$ were proposed. In the first case, a uniform probability distribution during the whole learning process was used. That is, for all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, K\}$,

$$\hat{\pi}_{k,A_k^{(n_k)}}(n) = \frac{1}{N_k}. \tag{5.43}$$

In the second case, at time interval n, higher probabilities are assigned to actions which have been played a smaller number of times during all time intervals between 0 and $n-1$. Let $T_{k,A_k^{(n_k)}}(n) \in \mathbb{N}$, with $k \in \mathcal{K}$ and $n_k \in \{1, \ldots, N_k\}$, be the number of times that player k has played action $A_k^{(n_k)}$ up to time interval n, i.e.,

$$T_{k,A_k^{(n_k)}}(n) = \sum_{s=0}^{n-1} \mathbb{1}_{\left\{a_k(s) = A_k^{(n_k)}\right\}}. \tag{5.44}$$

Then, the probability distribution to select the next action is the following:

$$\hat{\pi}_{k,A_k^{(n_k)}}(n) = \frac{\frac{1}{T_{k,A_k^{(n_k)}}(n)}}{\sum_{m=1}^{N_k} \frac{1}{T_{k,A_k^{(m)}}(n)}}, \tag{5.45}$$

where $T_{k,A_k^{(n_k)}}(0) = \delta$, with $\delta > 0$. We formalize the behavioral rule (5.42) in the Alg. 1, and we state its main property in the following proposition.

Algorithm 1 Learning the SE of the Game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ [Player $k \in \mathcal{K}$]

Require: At each instant $n > 0$: $\tilde{v}_k(n)$.

1: $n = 0$;

2: $\forall n_k \in \{1, \ldots, N_k\}$,

$$\begin{aligned} \hat{v}_{k,A_k^{(n_k)}}(0) &= 0, \\ \hat{\pi}_{k,A_k^{(n_k)}}(0) &= \frac{1}{N_k}. \end{aligned}$$

3: $a_k(0) \sim \hat{\pi}_k(0)$;

4: **for all** $n > 0$ **do**

5: $\quad \forall n_k \in \{1, \ldots, N_k\}$, update $\hat{\pi}_k(n)$.

6:

$$a_k(n) = \begin{cases} a_k(n-1) & \text{if} \quad \tilde{v}_k(n-1) = 1 \\ a_k(n) \sim \hat{\pi}_k(n) & \text{otherwise.} \end{cases}$$

end

Proposition 5.6.1 *The behavioral rule* (5.42) *converges to a SE of the game* $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ *in finite time if for all* $k \in \mathcal{K}$ *and for all* $n_k \in \{1, \ldots, N_k\}$, *it holds that,*

$$\hat{\pi}_{k,A_k^{(n_k)}}(n) > 0, \tag{5.46}$$

at each time interval $n \in \mathbb{N}$, *and assumptions* (i), (ii) *and* (iii) *always hold.*

The proof of Prop. 5.6.1 is as follows. Note that from (5.42), it can be concluded that if at a given time n, players play a SE action profile, then none of the players changes its own action in the following time intervals, i.e, the algorithm converges. Otherwise, unsatisfied players keep trying other actions. Now, from assumption (i), it is known that there always exists at least one SE. From assumption (ii) and (5.46), it can be concluded that all action profiles in the set $\mathcal{A}$ are played with a non-zero probability. Thus, since the set of actions is finite (assumption (iii)), one SE is certainly played during a finite number of time intervals.

From the reasoning above, it can be concluded that any probability distribution $\hat{\pi}_k(n)$ such that all actions have a non-zero probability of being played can be chosen as the probability distribution at time interval n for player k. However, as we shall see in the next section, the choice of this probability distributions might impact the convergence time.

5.6.3 Clipping Actions and SE

The behavioral rule (5.42) converges to a SE in pure strategies in finite time. However, condition (i) and (ii) are highly demanding. In real system scenarios, it is often observed that there might exists an action from a given player, which achieves satisfaction regardless of the actions adopted by all the other players. We refer to this kind of actions as *clipping actions* [85].

Definition 5.6.1 (Clipping Action) *In the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$, a player $k \in \mathcal{K}$ is said to have a clipping action $a_k \in \mathcal{A}_k$ if*

$$\forall \mathbf{a}_{-k} \in \mathcal{A}_{-k}, \quad a_k \in f_k\left(\mathbf{a}_{-k}\right). \tag{5.47}$$

As shown in the following proposition, the existence of clipping actions (assumption (ii) does not hold) in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ might inhibit the convergence of the behavioral rule in (5.42).

Proposition 5.6.2 *Consider the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ in satisfaction form. Assume the existence of at least one clipping action and denote it by $a_k^* \in \mathcal{A}_k$ for player k, with $k \in \mathcal{K}$. Then, if there exists a player $j \in \mathcal{K} \setminus \{k\}$, for which $f_j\left(a_k^*, \mathbf{a}_{-\{j,k\}}\right) = \emptyset$, $\forall \mathbf{a}_{-\{j,k\}} \in \prod_{i\in\mathcal{K}\setminus\{j,k\}} \mathcal{A}_i$. Then, the behavioral rule in* (5.42) *does not converge to a SE with strictly positive probability.*

The proof of Prop. 5.6.2 follows from the fact that at time $n > 0$ before convergence, the probability that player k plays the clipping action a_k^* is strictly positive (5.46). If player k plays a_k^*, by definition, there exist a player $j \neq k$ which would never be satisfied. Then, the behavioral rule does not converge to any SE.

5.7 Applications

In this section, we apply the concept of SE and ESE to the case of a classical interference channel [21] with 2 pairs of transmitter-receiver pairs sharing a common bandwidth. Here, the notions of SE and ESE are compared with the existing equilibrium notions such as NE and GNE. At the same time, the performance of the behavioral rules presented in Sec. 5.6 is evaluated in terms of convergence time to a satisfaction equilibrium.

5.7.1 QoS Provisioning in the Interference Channel

Consider a set $\mathcal{K} = \{1, 2\}$ of two transmitter-receiver pairs simultaneously operating over the same frequency band and thus, subject to mutual interference. Each transmitter communicates only with its corresponding receiver and any kind of message exchange aiming to achieve transmit cooperation is not considered. For all $(j, k) \in \mathcal{K}^2$, denote by $g_{j,k}$ and $p_k^{(n_k)}$ the channel gain between transmitter k and receiver j, and the n_k-th transmit power level of transmitter k, respectively. We denote by $\mathcal{A}_k = \left\{p_k^{(1)}, \ldots, p_k^{(N_k)}\right\}$, the set of all possible transmit power levels of player k. For all $k \in \mathcal{K}$, the minimum transmit power is $p_k^{(1)} = 0$ and the maximum transmit power is $p_k^{(N_k)} = p_{k,\max}$. The QoS metric, denoted by $u_k : \mathcal{A}_1 \times \mathcal{A}_2 \to \mathbb{R}_+$, of the transmitter-reciever pair k is its (Shannon) transmission rate in bits per second (bps). Thus, for all $(p_k, p_{-k}) \in \mathcal{A}_k \times \mathcal{A}_{-k}$, we write that,

$$u_k(p_k, p_{-k}) = \log_2\left(1 + \frac{p_k g_{k,k}}{\sigma_k^2 + \sum_{j \neq k}^{K} p_j g_{k,j}}\right) \text{ [bps/Hz]}. \tag{5.48}$$

Here, σ_k^2 is the noise level at receiver k and we denote the signal to noise ratio at the transmitter k by $\text{SNR}_k = \frac{p_{k,\max}}{\sigma_k^2}$. The QoS requirement for player k is to provide a transmission rate higher than Γ_k bps. Thus, we model the satisfaction correspondence f_k, as follows,

$$f_k(\mathbf{p}_{-k}) = \{p_k \in \mathcal{A}_k : u_k(p_k, p_{-k}) \geqslant \Gamma_k\}. \tag{5.49}$$

We assume also that transmitters associate different effort measures to each of their power levels. The higher the transmit power, the higher the effort.
This scenario is modeled by a game in classical normal form $\mathcal{G} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}\right\}$ and a game in satisfaction form $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$. We also model this scenario with a game $\hat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{u_k\}_{k\in\mathcal{K}}, \{f'_k\}_{k\in\mathcal{K}}\right\}$ in normal form with constrained action sets and a game in efficient satisfaction form $\tilde{\mathcal{G}} = \left(\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{c_k\}_{k\in\mathcal{K}}, \{f'_k\}_{k\in\mathcal{K}}\right)$, where, for all $k \in \mathcal{K}$, the cost or effort function c_k is defined as follows

$$c_k(p_k) = \begin{cases} p_{k,\max} + \delta & \text{if } \quad p_k = p_k^{(1)} \\ p_k & \text{if } \quad p_k \in \{p_k^{(2)}, \ldots, p_k^{(N_k)}\}, \end{cases} \tag{5.50}$$

where $\delta > 0$. Note that the most costly action is not to transmit. This choice is made to force the radio devices to transmit any time it is possible. The satisfaction correspondence f'_k is defined as follows:

$$f'_k(\mathbf{p}_{-k}) = \left\{p_k^{(1)}\right\} \cup \{p_k \in \mathcal{A}_k : u_k(p_k, p_{-k}) \geqslant \Gamma_k\}. \tag{5.51}$$

Here, we include the non-transmission action $p_k^{(1)} = 0$ in order to avoid an empty set of actions for players k, when there does not exist an action able to achieve the required minimum rate.

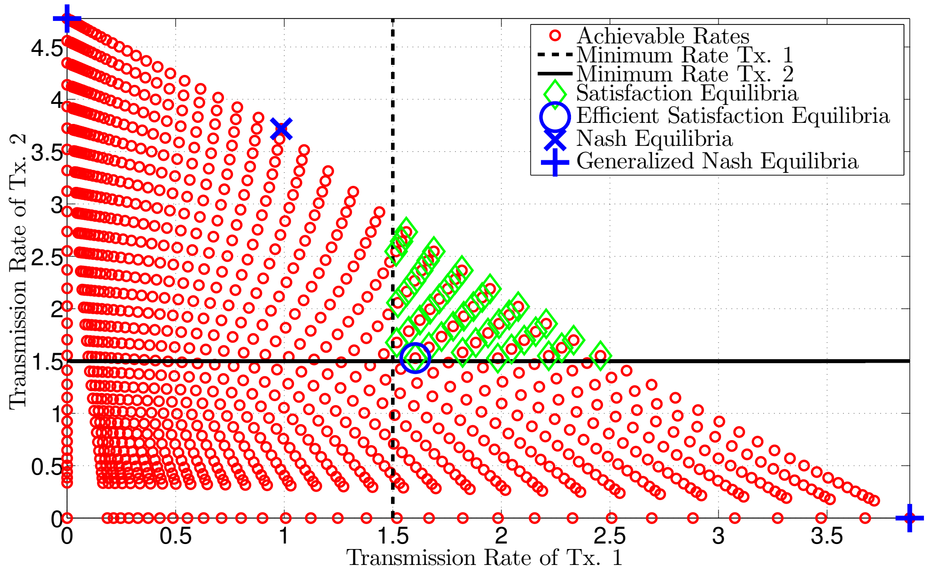

Figure 5.2: Achievable (Shannon) transmission rates $(u_1(p_1, p_2), u_2(p_1, p_2))$, for all $(p_1, p_2) \in \mathcal{A}_1 \times \mathcal{A}_2$, with SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels.

In Fig. 5.2, we plot (in red circles) all the achievable (Shannon) transmission rates for both transmitters, i.e., the pairs $(u_1(p_1, p_2), u_2(p_1, p_2))$, for all $(p_1, p_2) \in \mathcal{A}_1 \times \mathcal{A}_2$ and a particular channel realization. All the equilibria of the games $\mathcal{G}$, $\widehat{\mathcal{G}}$, $\hat{\mathcal{G}}$ and $\tilde{\mathcal{G}}$ are plotted. The unique NE of the game $\mathcal{G}$ is the action profile $(p_{1,\max}, p_{2,\max})$ (Def. 5.4.1). The game $\hat{\mathcal{G}}$ has two equilibria: the pairs $(0, p_{2,\max})$ and $(p_{1,\max}, 0)$ (Def. 5.4.2). The game $\widehat{\mathcal{G}}$ has multiple equilibria (Def. 5.2.1). In particular, note that in none of the equilibria of the games in classical normal form and normal form with constrained strategies, it is possible to simultaneously satisfy the QoS of both transmitters. In both cases, at most, only one transmitter can be satisfied. On the contrary, at the equilibrium of both games in satisfaction form and efficient satisfaction form, all players are able to satisfy their QoS demands. Importantly, the ESE satisfies the QoS condition for both transmitters with the lowest transmit power, while all the other SE require a higher transmission power. In particular, note that the set of GNE is not unitary while the set of ESE of $\tilde{\mathcal{G}}$ appear to be unitary. However, as shown before, the existence and uniqueness of the ESE and GNE are conditioned.

5.7.2 Clipping Actions in the Interference Channel

Note that the game $\widehat{\mathcal{G}}$ with the particular channel realization used in Fig. 5.2 possess at least one clipping action. For instance, when transmitter 2 transmits at the maximum power $p_{2,\max}$, it is always satisfied even if player 1 transmits at the maximum power (see the NE of the game $\mathcal{G}$ in Fig. 5.2). At the same time, if player 2 transmits at the maximum power, player 1 is unable to achieve satisfaction.

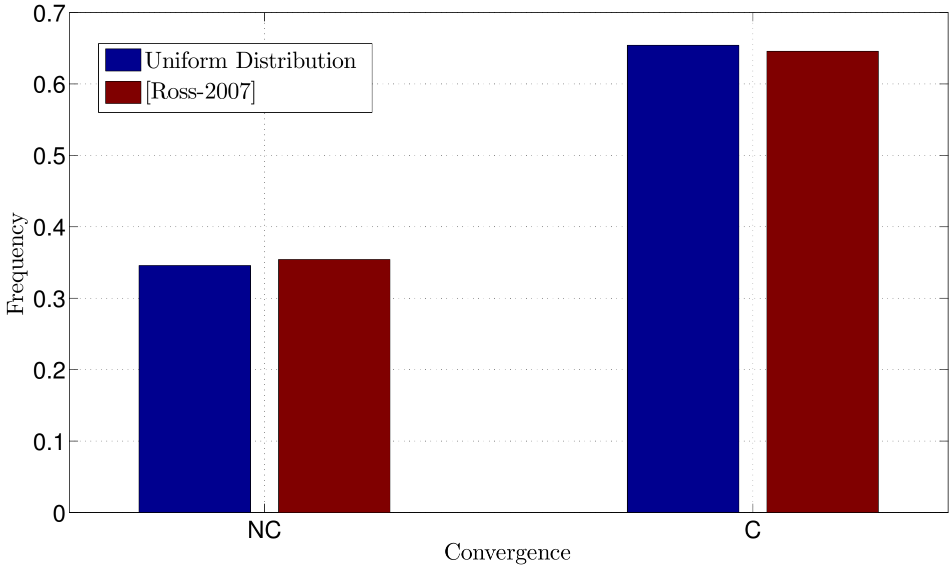

Figure 5.3: Histogram of the event of convergence or non-convergence of the learning algorithm (Alg. 1) in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$. Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} =$ 10 dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels.

Hence, if before observing convergence, transmitter 2 uses its maximum transmit power, then convergence to a SE is not observed neither in finite nor infinite time. In Fig. 5.3, we show an histogram of the convergence or not convergence of the algorithm. Here, we say that the algorithm does not convergence if during 100 consecutive time intervals, a given player does not change its current action, while the other still does (this implies that a clipping action might be being played). At each trial of the algorithm, we use the same channel realization used in Fig. 5.2. Note that independently of the probability distributions $\hat{\pi}_k(n)$ adopted by player k to try new actions, the event of one player playing a clipping strategy is non-negligible (0.3). In the particular case of the interference channel as treated here, the corresponding game is free of clipping actions if the simultaneous transmission at maximum power allows satisfaction. However, in this case, the distinction between SE and NE looses its importance since both equilibrium concepts would be able to give a satisfactory solution to the QoS problem. This observation leaves open the way for further research on learning algorithms in the context of the SE in the presence of clipping actions.

5.7.3 Convergence Time to the SE

Now, our interest focuses on the average time for converging to one SE of the game $\widehat{\mathcal{G}}$, when convergence is observed in the previous experiment. The convergence time is measured as the number of action updates required to each transmitter before convergence. In Fig. 5.4, we show an histogram of the convergence time when players try new actions with the probability distribution in (5.43) and (5.45). Note that in

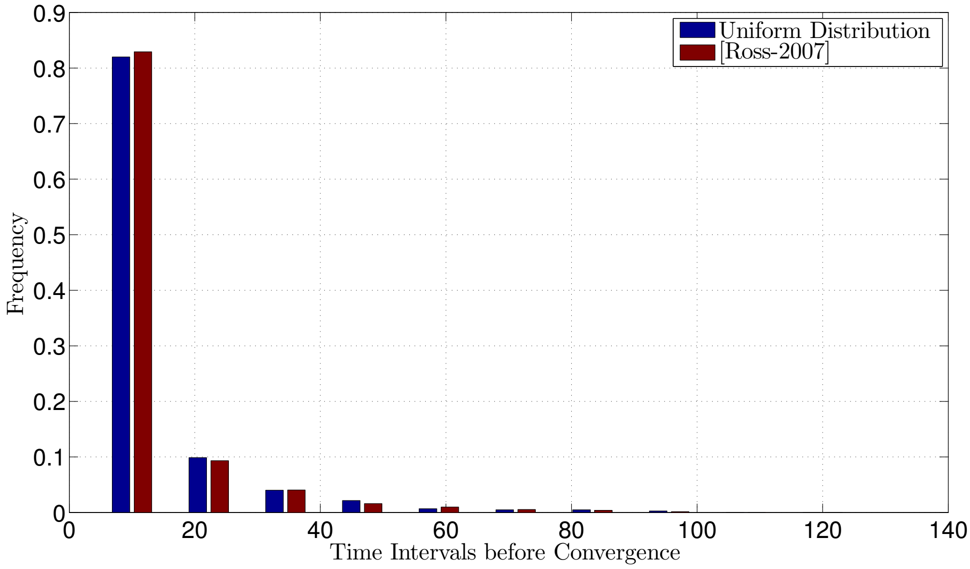

Figure 5.4: Histogram of the convergence time to a SE in the game $\widehat{\mathcal{G}} = \left\{\mathcal{K}, \{\mathcal{A}_k\}_{k\in\mathcal{K}}, \{f_k\}_{k\in\mathcal{K}}\right\}$ using the algorithm (Alg. 1). Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels.

this particular scenario, using a probability different from the uniform distribution does not bring a significant improvement. Interestingly, the histogram shows that if convergence is observed, most of the time (80%), satisfaction is achieved in less than 20 time intervals (action updates).
In Fig. 5.5, we plot the achieved transmission rate of both links at each instant n when the behavioral rule (5.42) is used. Therein, it can be observed that even though a transmitter is satisfied, and thus does not change its transmission power level, its instantaneous transmission rate changes due to the action updates of the other transmitters. Once both transmitters are satisfied, then, none of them changes its transmit powers.

5.8 Conclusions

The game formulation in satisfaction form (SF) and the notion of satisfaction equilibrium (SE) introduced in this chapter have been shown to be neatly adapted to model the problem of QoS provisioning in decentralized self-configuring networks. At the SE, all players are satisfied. On the contrary, when the QoS provisioning problem is modeled by games in classical normal form or normal form with constrained set of actions, equilibria where not all the players achieves satisfaction might be observed, even when there exist action profiles that allow the simultaneous satisfaction of all players. The notion of SE has been formalized in the context of pure and mixed strategies and its existence and uniqueness has been studied. In particular, when no SE exists neither in pure nor in mixed strategies, necessary and sufficient conditions

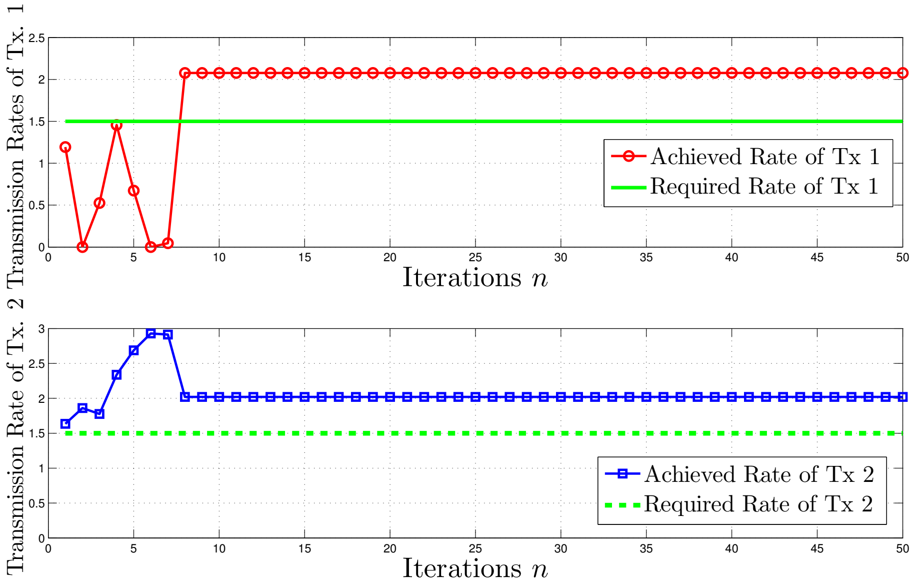

Figure 5.5: Instantaneous achieved rates of transmitter 1 (red) and 2 (blue).Here, SNR $= \frac{p_{k,\max}}{\sigma_k^2} = 10$ dBs, $(\Gamma_1, \Gamma_2) = (1.5, 1.5)$ bps and $N_1 = N_2 = 32$ power levels.

for the existence of an epsilon-SE has been presented. However, not all games in SF possess an ϵ-SE. Finally, a learning dynamics has been proposed to achieve SE. In particular, we remark that it requires only 1-bit feedback messages between the corresponding transmitter-receiver pairs. Nonetheless, the conditions for observing convergence to a pure SE are highly demanding and not practically appealing. This suggests that the design of algorithms such that at least one SE is learned in finite time and in a fully distributed fashion remains being an open problem.

Chapter 6

Conclusions and Perspectives

6.1 Conclusions

The contributions in this thesis can be classified in three main areas: (a) performance analysis and design of techniques for hierarchical spectrum access (HSA), (b) performance analysis and design of techniques for open spectrum access (OSA), and (c) mechanisms for quality of service (QoS) provisioning in both HSA and OSA.
In the context of HSA, the main contributions are the idea of opportunistic interference alignment (OIA) and bandwidth limiting (BL). The interest of the OIA scheme proposed in this thesis is that it provides a new concept of spectrum access opportunity (SAO). This novel idea of SAO corresponds to the unused spatial directions (SD) associated with the singular values of the channel matrix of a given transmitter-receiver primary pair using a water-filling power allocation (PA) scheme. In particular, this new type of SAO plays an important role in highly dense networks where classical SAOs are short-lasting and rare events. Interestingly, OIA has been shown to outperform the classical zero-forcing beamforming (ZFBF). The advantages of the OIA over the classical ZF techniques stem from the fact that it allows some interference to impair the primary signal space, while the ZF techniques avoid any interference in such space. In particular, the interference-free condition is not violated as long as the impaired dimensions of the primary signal space are those left over by the primary system due to limitations in the transmit power when it uses the water-filling power allocation. These additional degrees of freedom allow secondary systems to operate following a spectrum overlay policy even when all transmitters are equipped with the same number of antennas and the primary system achieves the highest achievable rate. This is in fact, the main advantage of OIA. Nonetheless, an important remark here is that for exploiting these new SAOs, the classical spectrum sensing does not suffice. Here, secondary systems must know, the exact channel matrix of the transmitter-receiver primary pair, which is possible only under certain cases. In the case of BL, the notion behind it is based on the idea of forcing the opportunistic systems to use a limited portion of the spectrum instead of all the available spectrum at a given time. By using an optimal BL, it is shown that it is possible to increase the network spectral efficiency. As shown in this thesis, depending on the topology of the network, there exists an optimal fraction

of spectrum that transmitters must use to maximize the network spectral efficiency. Such optimal fraction depends mainly on the number of active transmitters and its corresponding signal to noise ratios. The main drawback of this technique turns out to be the fact that, the calculation of the optimal spectrum fractions requires information that is not always available for each transmitter and thus, extra signaling might be required for its implementation.

In the context of OSA, the contribution presented in this thesis is basically the equilibrium analysis, in particular Nash equilibrium, of one of the simplest scenarios in spectrum sharing, i.e., the case where several transmitters communicate with the same receiver using a common set of frequency bands available for all the transmitters. This analysis is carried out considering that radio devices might either use several channels and use a water-filling power allocation scheme or use a unique channel at the maximum transmit power. Note that both the water-filling PA and the transmission at maximum power, in the corresponding scenarios, have a game-theoretic justification. Here, they correspond to the idea of best response. In both cases, the set of NE has been analyzed at it is shown that at least one NE always exists. The interest of the NE relies on the fact that once the network has achieved such a point, the transmit configuration of each radio device is optimal with respect to all the other radio devices' transmit configurations. In particular, in the case where the transmitters are limited to use only one channel, several NE are often observed, while in the other case, it is unique with probability one. More interestingly, it shown that the spectral efficiency observed by limiting the number of channels is better that the one observed when transmitters are left to use all the available channels, at least in average. This result is in line with the one described above for the case of HSA and suggests the paradoxical idea that in decentralized networks, reducing the amount of transmit configurations of radio devices, increases the network spectral efficiency. This result might appear limited due to the fact that it is strongly dependent on the topology of the network, however, it has been shown that the same effect is also observed in other topologies such as the classical interference channel. Another important remark is that an equilibrium, at least in the case of the multiple-transmitter unique-receiver scenario, can be achieved by simple dynamics such as best response dynamics and fictitious play. In fact, the best response can be implemented by a simple feedback message of the receiver containing the multiple access interference plus noise level on each of the frequency bands. On the contrary, fictitious play turns out to be more demanding in terms of required information.

The second contribution in OSA tackles precisely the design of techniques for achieving equilibrium independently of the topology of the network and with the minimum information each radio might possess about the network. Our main result in this scenario is a method to simultaneously estimate the average utility function achieved with each of its actions given the behavior of all the other players. This estimation is shown to be useful to calculate a soft-best response, which up to certain point can be interpreted as the best response in the sense of Nash. A soft-best response is a probability distribution that assigns high (resp. low) probabilities to the actions associated with high (resp. high) probability estimations. Here, we show that learning dynamics using such a soft-best response can be used to achieve performances which are close to equilibrium. At each learning step, such estimations are improved and at the same time used to tune the strategy of all radio devices. The advantage of the learning dynamics based on the utility estimation is that contrary to other learning dynamics, such as reinforcement learning, when convergence is observed, the converging point is an epsilon-close equilibrium. Another important remark is that this dynamics require only the knowledge of the achieved utility at each learning stage, which is very practically appealing. Interestingly, these dynamics have been shown to convergence in particular classes of games with a wide range of applications in self-configuring networks.

In the context of quality of service (QoS), the main contribution of this thesis is the formalization of the idea of satisfaction equilibrium (SE). Here, the existing idea of SE is put in terms of a fixed point inclusion, which facilitates the analysis of existence and uniqueness of the SE. As shown in this this thesis, classical concepts of equilibrium fail to properly model the QoS provisioning problem. For instance, it is shown that even when the QoS provisioning task is feasible, classical equilibrium concepts, such as GNE and NE, might lead to stable network states where some of the QoS requirement are not satisfied. On the contrary, the notion of SE has been shown to neatly model the QoS provisioning problem. Here, it is shown that at the SE, if it exists, any set of individual transmit configurations where the QoS requirement is satisfied is a stable point of the network. An interesting refinement of the SE, namely the efficient satisfaction equilibrium (ESE), is presented aiming to provide an equilibrium selection method. This refinement defines the concept of effort for satisfaction, for instance, the effort can be measured in terms of the transmit power. At the ESE, if it exists, all radio devices are able to satisfy their own QoS requirements by using the transmit configuration which requires the lowest effort. Another important contribution in this direction is a learning dynamics which allows radio devices to achieve SE using only one-bit feedback message at each stage of the learning process. Nonetheless, the convergence of this dynamics might be constrained depending on the network topology.

To conclude, we state that game theory has been shown to be a powerful tool to the analysis of DSCN. However, this theory might not be fully adapted to the scenarios encountered in wireless networks. Thus, some adaptations of the existing concepts are required. For instance, the idea of SE was motivated by real system needs encountered in DCSN. As in this case, many other concepts must be brought or adapted to the wireless communication domain to advance in the analysis and design of tools for optimal dynamic spectrum sharing in the context of DCSN.

6.2 Perspectives

6.2.1 On the Opportunistic Interference Alignment Strategy

Regarding the extensions of our opportunistic interference alignment scheme, we recall that our solution concerns only two MIMO links. The case where there exists several opportunistic devices and/or several primary devices remains to be studied in detail. Moreover, as recently shown by other authors, the OIA concept can be used in other network topologies, which broaden the applications of this technique. More importantly, some information assumptions could be relaxed to make the proposed approach more practical. This remark concerns CSI assumptions but also behavioral assumptions. In the first case, the lack of information in the secondary system can be tackled by using learning algorithms that allow cognitive radios to achieve IA using an iterative interacting process during a given period. However, this implies that the primary system must be tolerant to some amount of interference from the opportunistic system during the learning period. In the second case, it was assumed that the precoding scheme used by the primary transmitter is capacity-achieving, which allows the secondary transmitter to predict how the secondary transmitter is going to exploit its spatial resources. This behavioral assumption could be relaxed but some spatial sensing mechanisms should be designed to know which spatial modes can be effectively used by the secondary transmitter, which could be an interesting extension of the proposed scheme.

6.2.2 On the Bandwidth Limiting Strategy

Regarding the bandwidth limiting strategy, it has been shown that reducing the portion of spectrum that each radio device can use increases the spectral efficiency of the network. This result has been obtained considering certain topology and particular conditions on the channel statistics. Moreover, it has been assumed that the optimal fraction of spectrum is the same for all radio devices. The first extension of this work is then, to evaluate how the topology of the network influences this result. For instance, it must be determined if there exists a loss of optimality by imposing that all the radio devices must use the same fraction of spectrum. Here, the heterogeneity of the network might lead to a result where each radio device must use a different spectrum portion. Another important direction concerns the relaxation of the information required to implement it. A very interesting extension would be to design behavioral rules that allow radio devices to individually determine the optimal fraction of spectrum to be used with their local information.

6.2.3 On the Equilibrium Analysis of Spectrum Sharing Games

In this thesis, the interest on the Nash equilibria of DSCN relies on the fact that it allows to provide some prediction of its performance. However, we have shown that the NE analysis highly depends on the topology of the network. Here, a unified framework for the analysis of DCSN independently of the topology is missing. The importance of the research in this direction, relies on the fact that in DCSN, the topology is constantly changing. Thus, an analysis such as the one presented in this thesis, is limited to the time the channels remain constant and the topology remains unchanged. Ideally, a general framework for the analysis of DSCN must take into consideration these facts. Indeed, particular attention must be given to the fact that radio devices operate during the time their need of communication exists, that is, the number of active transmitters is constantly changing, all the elements of the network might be constantly moving, etc. Another direction for further research is to consider that not all the radio devices are equally reliable. For instance, the fact that there might exist radio devices aiming to break off the communications, either because it is in its own benefit or simply because external elements are designed to attack the network under study. This malicious behavior has not been taken into account in this analysis.

6.2.4 On the Schemes for Learning Equilibria

Learning equilibrium in DCSN is one of the most interesting lines for further research. As we mentioned before, the different concepts of equilibrium, for instance Nash equilibrium, correlated equilibrium, satisfaction equilibrium, provide an estimation of the performance of the network. However, achieving such equilibrium in a fully decentralized network remains an open issue. As shown in this thesis, in some network topologies, very simple behavioral rules, e.g., best response dynamics or fictitious play, lead to equilibrium. However, the convergence or non-convergence of these algorithms is strongly dependent on the topology. For instance, when radio devices aim to maximize their Shannon rates, both BRD and FP converge to an NE in the parallel multiple access channel, while in the case of the parallel interference channel, such a convergence is not ensured. In general, a behavioral rule that allows all radio devices to achieve NE in finite time, independently of the topology, does not exist. Often, algorithms achieve epsilon-close equilibrium performance after certain number of iterations. However, depending on the application, the global performance of the network might be highly sensitive to this learning time. Thus, the design of algorithms that allow DCSN to achieve equilibrium performance in a short time with minimum feedback remains an open problem.

6.2.5 On the Quality of Service Provisioning

In this thesis, the concept of SE has been formalized and has been shown to be particularly suited to model the problem of QoS provisioning. In this direction, only few applications have been presented. However, many theoretical aspects remain to be completed. For instance, exploiting the formulation as a fixed point inclusion of the SE to obtain more general results on the existence or uniqueness of the SE and ESE is one interesting research direction. More interesting, a generalization of the SE and ESE concept to dynamic games e.g., stochastic games, remains to be formulated. This formulation in dynamic games would allow us to model the time varying nature of wireless communications networks. From a practical point of view, general algorithms for achieving SE and ESE in a fully decentralized fashion remain also to be designed. Here, a particular class of actions has been identified to be very complicated to deal with. This is the case of the clipping strategies, which at a given point of time can make standard algorithms not to converge even when the existence of at least one SE is ensured. The design of behavioral rules such that SE, ESE or at least epsilon-SE can be achieved in the presence of clipping strategies remains being an open issue and so far, the main constraint on the application of SE to the spectrum sharing games.

Bibliography

[1] E. Altman, N. Bonneau, and M. Debbah, "Correlated equilibrium in access control for wireless communications," *Lecture Notes in Computer Science*, pp. 173–183, 2006.

[2] E. Altman and E. Solan, "Constrained Games: The Impact of the Attitude to Adversarys Constraints," *IEEE Transactions on Automatic Control*, vol. 54, no. 10, pp. 2435–2440, 2009.

[3] R. Aumann, "Subjectivity and correlation in randomized strategies," *Journal of Mathematical Economics*, vol. 1, no. 01, pp. 67–96, 1974.

[4] R. G. Bartle, *The Elements of Integration and Lebesgue Measure.* Wiley-Interscience, 1995.

[5] E. V. Belmega, M. Jungers, and S. Lasaulce, "A generalization of a trace inequality for positive definite matrices," *The Australian Journal of Mathematical Analysis and Applications (AJMAA)*, 2010.

[6] E. V. Belmega, S. Lasaulce, and M. Debbah, "Power allocation games for MIMO multiple access channels with coordination," *IEEE Trans. on Wireless Communications*, vol. 8, no. 6, pp. 3182–3192, June 2009.

[7] E. V. Belmega, H. Tembine, and S. Lasaulce, "Learning to precode in outage minimization games over MIMO interference channels," in *IEEE Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2010, pp. 1–5.

[8] E. Belmega, S. Lasaulce, and M. Debbah, "Decentralized handovers in cellular networks with cognitive terminals," *3rd Intl. Symp. on Communications, Control and Signal Processing - ISCCSP*, March 2008.

[9] M. Benaïm and M. W. Hirsch, "Mixed equilibria and dynamical systems arising from fictitious play in perturbed games," *Games and Economic Behavior*, vol. 29, no. 1–2, pp. 36–72, 1999.

[10] M. Bennis and S. M. Perlaza, "Decentralized cross-tier interference mitigation in cognitive femtocell networks," in *IEEE Intl. Conference on Communications (ICC2011)*, Kyoto, Japan, June 2011.

[11] K. C. Border, *Fixed Point Theorems with Applications to Economics and Game Theory.* New York, NY, USA: Cambridge University Press, 1985.

[12] V. Borkar, "Stochastic approximation with two timescales," *Systems Control Lett.*, vol. 29, pp. 291–294, 1997.

[13] ——, "Reinforcement learning in Markovian evolutionary games," *Advances in Complex Systems (ACS)*, vol. 05, no. 01, pp. 55–72, 2002.

[14] ——, "Sthocastic approximation: A dynamical system viewpoint," *Cambdrige University Press*, 2008.

[15] D. Braess, "*Über* ein Paradoxon aus der Verkehrsplanung," *Unternehmensforschung*, vol. 24, no. 5, pp. 258 – 268, May 1969.

[16] G. W. Brown, "Iterative solution of games by fictitious play," *Activity Analysis of Production and Allocation*, vol. 13, no. 1, pp. 374–376, 1951.

[17] M. Buddhikot, "Understanding dynamic spectrum access: Models, taxonomy and challenges," in *IEEE DySPAN*, April 2007.

[18] M. F. Bush R., "Stochastic models of learning," *Wiley Sons, New York.*, 1955.

[19] V. Cadambe, S. Jafar, and S. Shamai, "Interference alignment on the deterministic channel and application to fully connected AWGN interference networks," in *Proc. IEEE Information Theory Workshop (ITW)*, Porto, Portugal, May. 2008.

[20] V. Cadambe and S. Jafar, "Interference alignment and degrees of freedom of the k-user interference channel," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[21] A. Carleial, "Interference channels," *IEEE Trans. Inform. Theory*, vol. 24, no. 1, pp. 60–70, 1978.

[22] C. E. Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inform. Theory*, vol. 48, pp. 637–650, Mar. 2002.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Wiley-Interscience, 1991.

[24] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 10, pp. 886–893, October 1952.

[25] J. Dumont, W. Hachem, S. Lasaulce, P. Loubaton, and J. Najim, "On the capacity achieving covariance matrix of Rician MIMO channels: An asymptotic approach," *IEEE Trans. on Info. Theory*, vol. 56, no. 3, pp. 1048–1069, March 2010.

[26] Émile Borel, “La théorie du jeu et les équations à noyau symétrique,” *Comptes Rendus de l'Académie des Sciences*, vol. 173, pp. 1304–1308, Sept. 1921.

[27] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” *IEEE Journal on Selected Areas on Communications, Special issue on adaptive, Spectrum Agile and Cognitive Wireless Networks*, vol. 25, no. 3, pp. 517–528, 2007.

[28] K. Fan, “Fixed-point and Minimax Theorems in Locally Convex Topological Linear Spaces,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 2, pp. 121–126, 1952.

[29] N. Fawaz, K. Zarifi, M. Debbah, and D. Gesbert, “Asymptotic capacity and optimal precoding strategy of multi-level precode & forward in correlated channels,” *IEEE Trans. on Inform. Theory*, vol. 57, no. 4, p. 2050, Apr. 2011.

[30] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” Bell Labs, Tech. Rep., 1996.

[31] D. Fudenberg and J. Tirole, “Game theory,” *MIT Press*, 1991.

[32] K. Gomadam, V. Cadambe, and S. Jafar, “Approaching the capacity of wireless networks through distributed interference alignment,” in *Proc. IEEE Global Telecommunications Conference (GLOBECOM)*, New Orleans, USA, Dec. 2008.

[33] M. Haddad, A. Hayar, and M. Debbah, “Spectral efficiency of spectrum-pooling systems,” *IET Communications*, vol. 2, no. 6, pp. 733–741, July 2008.

[34] ——, “Spectral efficiency of cognitive radio systems,” *IEEE Global Telecommunications Conference*, pp. 4165–4169, Nov. 2007.

[35] S. Hart and A. Mas-Colell, “A simple adaptive procedure leading to correlated equilibrium,” *Econometrica*, vol. 68, no. 5, pp. 1127–1150, September 2000.

[36] S. Haykin, “Cognitive radio: Brain-empowered wireless communications,” *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[37] J. Hofbauer and E. Hopkins, “Learning in perturbed asymmetric games,” Edinburgh School of Economics, University of Edinburgh, ESE Discussion Papers 53, Apr. 2004.

[38] S. Jafar and M. Fakhereddin, “Degrees of freedom for the MIMO interference channel,” *IEEE Trans. Inform. Theory*, vol. 53, no. 7, pp. 2637–2642, July 2007.

[39] J. S. Jordan, “Three problems in learning mixed-strategy Nash equilibria,” *Games and Economic Behavior*, vol. 5, no. 3, pp. 368–386, 1993.

[40] S. Kakutani, "A generalization of Brouwer's fixed point theorem," *Duke Mathematical Journal*, vol. 8, pp. 457–459, 1941.

[41] Y. M. Kaniovski and H. P. Young, "Learning dynamics in games with stochastic perturbations," *Games and Economic Behavior*, vol. 11, no. 2, pp. 330–363, November 1995.

[42] B. Knaster and A. Tarski, "Un théorème sur les fonctions d'ensembles," *Ann. Soc. Polon. Math.*, vol. 6, pp. 133–134, 1928.

[43] V. R. Konda, John, and J. N. Tsitsiklis, "Actor-critic algorithms," in *SIAM Journal on Control and Optimization*. MIT Press, 2001, pp. 1008–1014.

[44] V. R. Konda and V. Borkar, "Actor-critic–type learning algorithms for Markov decision processes," *SIAM J. Control Optim.*, vol. 38, no. 1, pp. 94–123, 1999.

[45] I. Krikidis, "Space alignment for cognitive transmission in MIMO uplink channels," *EURASIP J. Wireless Comm. and Networking*, vol. 2010, 2010.

[46] S. Lasaulce, M. Debbah, and E. Altman, "Methodologies for analyzing equilibria in wireless games," *IEEE Signal Processing Magazine, Special issue on Game Theory for Signal Processing*, vol. 26, no. 5, pp. 41–52, Sep. 2009.

[47] S. Lasaulce and H. Tembine, *Game Theory and Learning in Wireless Networks: Fundamentals and Applications.* Elsevier Academic Press, 2011.

[48] S. D. Leslie and E. J. Collins, "Convergent multiple-timescales reinforcement learning algorithms in normal form games," *Ann. Appl. Probab.*, vol. 13, no. 4, pp. 1231–1251, 2003.

[49] R. D. Luce, *Individual Choice Behavior: A Theoretical Analysis.* New York: Wiley, 1959.

[50] A. B. Mackenzie and L. Da Silva, *Game Theory for Wireless Engineers (Synthesis Lectures on Communications)*, 1st ed. Morgan & Claypool Publishers, May 2006.

[51] A. B. Mackenzie and S. B. Wicker, "Game theory and the design of self-configuring, adaptive wireless networks," *IEEE Communications Magazine*, vol. 39, no. 11, pp. 126–131, 2001.

[52] M. Maddah-Ali, A. Motahari, and A. Khandani, "UW-ECE-2006-12 - communication over X channel: Signalling and multiplexing gain," University of Waterloo, Tech. Rep., 2006.

[53] ——, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inform. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[54] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications.* Academic Press, 1979.

[55] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457–483, 1967.

[56] D. McFadden, "Quantal choice analysis: A survey," *Annals of Economic and Social Measurement*, vol. 5, no. 4, pp. 363–390, June 1976.

[57] R. D. McKelvey and T. R. Palfrey, "Quantal response equilibria for normal form games," *Games and Economic Behavior*, vol. 10, no. 1, pp. 6 – 38, 1995.

[58] ——, "Quantal response equilibria for extensive form games," *Experimental Economics*, vol. 1, no. 1, pp. 9–41, June 1998.

[59] P. Mertikopolous, E. V. Belmega, A. Moustakas, and S. Lasaulce, "Dynamic power allocation in parallel multiple access channels," in *5th International ICST Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS)*, Paris, France, May 2011.

[60] F. Meshkati, M. Chiang, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient power control in multi-carrier CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 6, pp. 1115–1129, 2006.

[61] F. Meshkati, H. V. Poor, S. C. Schwartz, and N. B. Mandayam, "An energy-efficient approach to power control and receiver design in wireless data networks," *IEEE Transactions on Communications*, vol. 53, no. 11, pp. 1885–1894, 2005.

[62] D. Minoli and E. Minoli, *Delivering voice over IP networks.* New York, NY, USA: John Wiley & Sons, Inc., 1998.

[63] D. Monderer and L. S. Shapley, "Fictitious play property for games with identical interests," *Int. J. Economic Theory*, vol. 68, pp. 258–265, 1996.

[64] ——, "Potential games," *Games and Economic Behavior*, vol. 14, pp. 124–143, 1996.

[65] R. Müller, "On the asymptotic eigenvalue distribution of concatenated vector-valued fading channels," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 2086–2091, Jul. 2002.

[66] J. F. Nash, "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 36, no. 1, pp. 48–49, 1950.

[67] F. Neeser and J. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1293–1302, Jul. 1993.

[68] J. V. Neumann and O. Morgenstern, "Theory of games and economic behavior," *Princeton University Press*, 1944.

[69] A. Neyman, "Correlated equilibrium and potential games," *International Journal of Game Theory*, vol. 26, no. 2, pp. 223–227, 1997.

[70] A. Neyman and S. Sorin, *Stochastic Games and Applications.* NATO Science Series, 1999.

[71] J.-S. Pang, G. Scutari, F. Facchinei, and C. Wang, "Distributed power allocation with rate constraints in Gaussian parallel interference channels," *IEEE Trans. on Info. Theory*, vol. 54, no. 8, pp. 3471–3489, Aug. 2008.

[72] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications.* Cambridge Univ. Press, 2003.

[73] S. M. Perlaza, E. V. Belmega, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks," in *3rd ICST/ACM International Workshop on Game Theory in Communication Networks*, Pisa, Italy, Oct. 2009.

[74] S. M. Perlaza and M. Debbah, "Modeling noisy feedback in decentralized self-configuring networks," in *45th Annual Asilomar Conference on Signals, Systems, and Computers (Asilomar2011)*, Asilomar, Pacific Grove, 2011.

[75] S. M. Perlaza, M. Debbah, S. Lasaulce, and H. Bogucka, "On the benefits of bandwidth limiting in decentralized vector multiple access channels," in *4th Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Comm. (CROWNCOM)*, Hanover, Germany, May 2009.

[76] S. M. Perlaza, M. Debbah, S. Lasaulce, and J.-M. Chaufray, "Opportunistic interference alignment in MIMO interference channels," in *Proc. IEEE 19th Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Cannes, France, Sept. 2008.

[77] S. M. Perlaza, N. Fawaz, M. Debbah, and S. Lasaulce, "Alignement d'interférence opportuniste avec des terminaux multi-antennes," in *Gretsi Conference*, Dijon, France, 2009.

[78] S. M. Perlaza, N. Fawaz, S. Lasaulce, and M. Debbah, "From spectrum pooling to space pooling: Opportunistic interference alignment in MIMO cognitive networks," *IEEE Trans. in Signal Processing*, vol. 58, no. 7, pp. 3728–3741, July 2010.

[79] S. M. Perlaza, S. Lasaulce, and M. Debbah, "Game theory for dynamic spectrum sharing," *Cognitive Radio Networks: Architectures, Protocols and Standards*, 2008.

[80] ——, "Equilibria of channel selection games in parallel multiple access channels," *EURASIP Journal in Wireless Communications (Submitted)*, Dec. 2011.

[81] S. M. Perlaza, S. Lasaulce, H. Tembine, and M. Debbah, "Radio resource sharing in decentralized wireless networks: A logit equilibrium approach," *Submitted to IEEE Trans. on Signal Processing.*, Jan. 2011.

[82] S. M. Perlaza, H. Tembine, and S. Lasaulce, "How can ignorant but patient cognitive terminals learn their strategy and utility?" in *the 11th IEEE Intl. Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2010)*, Marrakech, Morocco, June 2010.

[83] S. M. Perlaza, H. Tembine, S. Lasaulce, and M. Debbah, "Satisfaction equilibrium: A general framework for QoS provisioning in self-configuring networks," in *the IEEE Global Communications Conference (GLOBECOM)*, Miami, USA, Dec. 2010.

[84] ——, "Learning to use the spectrum in self-configuring heterogeneous networks: A logit equilibrium approach," in *4th International ICST Workshop on Game Theory in Communication Networks*, 2011.

[85] ——, "A general framework for Quality-Of-Service Provisioning in decentralized networks," *submitted to the IEEE Journal in Selected Topics in Signal Processing. Special Issue in Game Theory for Signal Processing.*, vol. 6, no. 2, pp. 1–13, Apr. 2012.

[86] S. M. Perlaza, H. Tembine, S. Lasaulce, and V. Quintero-Florez, "On the fictitious play and channel selection games," in *IEEE Latin-American Conference on Communications (LATINCOM)*, Bogota, Colombia, Sept. 2010, pp. 1–5.

[87] S. M. Perlaza, M. Debbah, S. Lasaulce, and J.-M. Chaufray, "Opportunistic interference alignment in mimo interference channels," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2008.

[88] S. W. Peters and R. W. Heath, "Interference alignment via alternating minimization," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009.

[89] A. Poznyak and K. Najim, "Learning through reinforcement for N-person repeated constrained games," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 32, no. 6, pp. 759 – 771, dec. 2002.

[90] L. Rose, S. M. Perlaza, and M. Debbah, "On the Nash equilibria in decentralized parallel interference channels," in *IEEE Workshop on Game Theory and Resource Allocation for 4G*, Kyoto, Japan, Jun. 2011.

[91] L. Rose, S. M. Perlaza, S. Lasaulce, and M. Debbah, "Learning equilibria with partial information in wireless networks," *IEEE Communication Magazine. Special Issue on Game Theory for Wireless Communications*, Aug. 2011.

[92] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[93] S. Ross and B. Chaib-draa, "Satisfaction equilibrium : Achieving cooperation in incomplete information games," in *the 19th Canadian Conf. on Artificial Intelligence*, 2006.

[94] ——, "Learning to play a satisfaction equilibrium." in *Workshop on Evolutionary Models of Collaboration*, 2007.

[95] P. Sastry, V. Phansalkar, and M. Thathachar, "Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 5, pp. 769–777, May 1994.

[96] G. Scutari, S. Barbarossa, and D. Palomar, "Potential games: A framework for vector power control problems with coupled constraints," *in Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 4, May 2006.

[97] ——, "Potential games: A framework for vector power control problems with coupled constraints," *Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2006.

[98] G. Scutari, D. Palomar, and S. Barbarossa, "Optimal linear precoding strategies for wideband non-cooperative systems based on game theory – part II: Algorithms," *IEEE Trans. on Signal Processing*, vol. 56, no. 3, pp. 1250–1267, mar. 2008.

[99] ——, "Optimal linear precoding strategies for wideband noncooperative systems based on game theory – Part I: Nash equilibria," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1230 –1249, mar. 2008.

[100] ——, "The MIMO iterative waterfilling algorithm," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[101] ——, "Simultaneous iterative water-filling for gaussian frequency-selective interference channels," in *IEEE Intl. Symp. on Information Theory*, July 2006.

[102] X. Shang, B. Chen, and M. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4313–4320, Sep 2006.

[103] L. S. Shapley, "Some topics in two-person games," *Advances in Game Theory, M. Dresher, Lloyd S. Shapley and A. W. Tucker, eds., Princeton University Press.*, pp. 1–28, 1964.

[104] J. W. Silverstein and Z. D. Bai, "On the empirical distribution of eigenvalues of a class of large dimensional random matrices," *J. Multivariate. Anal.*, vol. 54, no. 2, pp. 175–192, 1995.

[105] E. Solan, "Stochastic games," *Encyclopedia of Database Systems, Springer*, 2009.

[106] V. Srivastava, J. Neel, A. B. Mackenzie, R. Menon, L. A. Dasilva, J. E. Hicks, J. H. Reed, and R. P. Gilles, "Using game theory to analyze wireless ad hoc networks," *Communications Surveys & Tutorials, IEEE*, vol. 7, no. 4, pp. 46–56, 2005.

[107] R. S. Sutton, D. Mcallester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," in *In Advances in Neural Information Processing Systems 12*, vol. 12, 2000, pp. 1057–1063.

[108] E. Telatar, "Capacity of multi-antenna Gaussian channels," Bell Labs., Tech. Rep., 1995.

[109] ——, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–596, 1999.

[110] J. Thukral and H. Bölcskei, "Interference alignment with limited feedback," in *Proc. IEEE Intl. Symp. on Information Theory (ISIT)*, Seoul, Korea, Jun. 2009.

[111] R. Tresch, M. Guillaud, and E. Riegler, "On the achievability of interference alignment in the K-user constant MIMO interference channel," in *Proc. IEEE Workshop on Statistical Signal Processing (SSP)*, Cardiff, Wales, UK, Sept. 2009.

[112] Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications.* Now Publishers Inc., 2004.

[113] T. Ui, "Discrete concavity for potential games," *International Game Theory Review (IGTR)*, vol. 10, no. 01, pp. 137–143, 2008.

[114] S. Vishwanath and S. Jafar, "On the capacity of vector Gaussian interference channels," in *Proc. IEEE Inform. Theory Workshop (ITW)*, San Antonio, USA, Oct 2004.

[115] M. Voorneveld, "Best-response potential games," *Economics Letters*, vol. 66, no. 3, pp. 289–295, March 2000.

[116] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel," in *Proc. Annual Information Theory and Applications Workshop*, San Diego, CA., Jan. 2007.

[117] R. Wilson, "Computing equilibria of N-person games," *SIAM Journal on Applied Mathematics*, vol. 21, no. 1, pp. 80–87, Jul. 1971.

[118] Y. Xing and R. Chandramouli, "Stochastic learning solution for distributed discrete power control game in wireless data networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 4, pp. 932–944, 2008.

[119] H. P. Young, "The evolution of conventions," *Econometrica*, vol. 61, no. 1, pp. 57–84, 1993.

[120] ——, "Strategic learning and its limits (Arne Ryde memorial lectures sereis)," *Oxford University Press, USA*, 2004.

[121] W. Yu, W. Rhee, S. Boyd, and J. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. on Info. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.

[122] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.

[123] W. Zhong, Y. Xu, M. Tao, and Y. Cai, "Game theoretic multimode precoding strategy selection for MIMO multiple access channels," *IEEE Signal Processing Letters*, vol. 17, no. 6, pp. 563 – 566, jun. 2010.

Appendix A

Definitions

In this appendix, we present useful definitions and previous results used in the proofs of Appendix C.

Definition A.0.1 *Let* $\mathbf{X}$ *be an* $n \times n$ *random matrix with empirical eigenvalue distribution function* $F_X^{(n)}$*. We define the following transforms associated with the distribution* $F_X^{(n)}$*, for* $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$*:*

$$\text{Stieltjes transform: } G_X(z) \triangleq \int_{-\infty}^{\infty} \frac{1}{t-z} \mathrm{d}F_X^{(n)}(t), \tag{A.0.1}$$

$$\Upsilon_X(z) \triangleq \int_{-\infty}^{\infty} \frac{zt}{1-zt} \mathrm{d}F_X^{(n)}(t), \tag{A.0.2}$$

$$\text{S-transform: } S_X(z) \triangleq \frac{1+z}{z} \Upsilon_X^{-1}(z), \tag{A.0.3}$$

where the function $\Upsilon_X^{-1}(z)$ *is the reciprocal function of* $\Upsilon_X(z)$*, i.e.,*

$$\Upsilon_X^{-1}(\Upsilon_X(z)) = \Upsilon_X(\Upsilon_X^{-1}(z)) = z. \tag{A.0.4}$$

From (A.0.1) and (A.0.2), we obtain the following relationship between the function $\Upsilon_X(z)$ (named Υ-transform in [112]) and the Stieltjes transform $G_X(z)$,

$$\Upsilon_X(z) = -1 - \frac{1}{z} G_X\left(\frac{1}{z}\right). \tag{A.0.5}$$

Appendix B

Proof of Lemma 2.1.4

Here, we prove *Lemma 2.1.4* which states that: if a matrix $\mathbf{V}_2$ satisfies the condition $\tilde{\mathbf{H}}_1\mathbf{V}_2 = \mathbf{0}_{(N_1-S)\times L_2}$ then it meets the IA condition (2.1.3).
Proof: Let $\mathbf{H}_{11} = \mathbf{U}_{H_{11}}\mathbf{\Lambda}_{H_{11}}\mathbf{V}_{H_{11}}^H$ be a sorted SVD of matrix $\mathbf{H_{11}}$, with $\mathbf{U}_{H_{11}}$ and $\mathbf{V}_{H_{11}}$, two unitary matrices of sizes $N_1 \times N_1$ and $M_1 \times M_1$, respectively, and $\mathbf{\Lambda}_{H_{11}}$ an $N_1 \times M_1$ matrix with main diagonal $\left(\lambda_{H_{11},1}, \ldots, \lambda_{H_{11},\min(N_1,M_1)}\right)$ and zeros on its off-diagonal, such that $\lambda^2_{H_{11},1} \geqslant \lambda^2_{H_{11},2} \geqslant \ldots \geqslant \lambda^2_{H_{11},\min(N_1,M_1)}$. Given that the singular values of the matrix $\mathbf{H}_{11}$ are sorted, we can write matrix $\mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H$ as a block matrix,

$$\mathbf{\Lambda}_{H_{11}}\mathbf{P}_1\mathbf{\Lambda}_{H_{11}}^H = \begin{pmatrix} \mathbf{\Psi} & \mathbf{0}_{m_1\times(N_1-m_1)} \\ \mathbf{0}_{(N_1-m_1)\times m_1} & \mathbf{0}_{(N_1-m_1)\times(N_1-m_1)} \end{pmatrix}, \tag{B.0.1}$$

where the diagonal matrix $\mathbf{\Psi}$ of size $m_1\times m_1$ is $\mathbf{\Psi} = \operatorname{diag}\left(\lambda^2_{H_{11},1}\, p_{1,1}, \ldots, \lambda^2_{H_{11},m_1}\, p_{1,m_1}\right)$. Now let us split the interference-plus-noise covariance matrix (2.11) as:

$$\mathbf{R} = \begin{array}{c} \\ m_1 \updownarrow \\ N_1-m_1 \updownarrow \end{array} \begin{array}{c} \begin{array}{cc} \overset{m_1}{\longleftrightarrow} & \overset{N_1-m_1}{\longleftrightarrow} \end{array} \\ \begin{pmatrix} \mathbf{R}_1+\sigma_1^2\mathbf{I}_{m_1} & \mathbf{R}_2 \\ \mathbf{R}_2^H & \mathbf{R}_3+\sigma_1^2\mathbf{I}_{N_1-m_1} \end{pmatrix} \end{array}, \tag{B.0.2}$$

where $(\mathbf{R}_1 + \sigma_1^2\mathbf{I}_{m_1})$ and $(\mathbf{R}_3 + \sigma_1^2\mathbf{I}_{N_1-m_1})$ are invertible Hermitian matrices, and matrices $\mathbf{R}_1$, $\mathbf{R}_2$ and $\mathbf{R}_3$ are defined from (2.11) and (2.13) as

$$\mathbf{R}_1 \triangleq \tilde{\mathbf{H}}_1\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\tilde{\mathbf{H}}_1^H, \tag{B.0.3}$$
$$\mathbf{R}_2 \triangleq \tilde{\mathbf{H}}_1\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\tilde{\mathbf{H}}_2^H, \tag{B.0.4}$$
$$\mathbf{R}_3 \triangleq \tilde{\mathbf{H}}_2\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\tilde{\mathbf{H}}_2^H. \tag{B.0.5}$$

Now, by plugging expressions (B.0.1) and (B.0.2) in (2.12), the IA condition can be rewritten as follows:

$$\begin{aligned} &\log_2\left|\sigma_1^2\mathbf{I}_{m_1}+\mathbf{\Psi}\right| - \log_2\left|\sigma_1^2\mathbf{I}_{N_1}\right| = \log_2\left|\mathbf{R}_1+\sigma_1^2\mathbf{I}_{m_1}+\mathbf{\Psi}\right| \\ &- \log_2\left|\mathbf{R}_1+\sigma_1^2\mathbf{I}_{m_1}\right| - \\ &\log_2\left(\frac{\left|\mathbf{R}_3+\sigma_1^2\mathbf{I}_{N_1-m_1}-\mathbf{R}_2^H\left(\mathbf{R}_1+\sigma_1^2\mathbf{I}_{m_1}\right)^{-1}\mathbf{R}_2\right|}{\left|\mathbf{R}_3+\sigma_1^2\mathbf{I}_{N_1-m_1}-\mathbf{R}_2^H\left(\mathbf{R}_1+\sigma_1^2\mathbf{I}_{m_1}+\mathbf{\Psi}\right)^{-1}\mathbf{R}_2\right|}\right). \end{aligned} \tag{B.0.6}$$

Note that there exists several choices for the submatrices $\mathbf{R}_1$, $\mathbf{R}_2$, and $\mathbf{R}_3$ allowing the equality in (B.0.6) to be met. We see that a possible choice in order to meet the IA condition is $\mathbf{R}_1 = \mathbf{0}$, $\mathbf{R}_2 = \mathbf{0}$, independently of the matrix $\mathbf{R}_3$. Thus, from (B.0.3) and (B.0.4) we have $\mathbf{R}_1 = \mathbf{0}$ and $\mathbf{R}_2 = \mathbf{0}$ by imposing the condition $\tilde{\mathbf{H}}_1\mathbf{V}_2 = \mathbf{0}_{m_1 \times L_2}$, for any given PA matrix $\mathbf{P}_2$, which concludes the proof. $\square$

Appendix C

Proof of Proposition 2.1.1

In this appendix, we provide a proof of Prop. 2.1.1 on the asymptotic expression of the opportunistic transmission rate per antenna, defined by

$$\bar{R}_{2,\infty}(\mathbf{P}_2,\sigma^2) \triangleq \lim_{\substack{\forall(i,j)\in\{1,2\}^2,\, N_i, M_j\to\infty \\ \forall(i,j)\in\{1,2\}^2,\, \frac{M_j}{N_i}\to\alpha_{ij}<\infty}} \bar{R}_2(\mathbf{P}_2,\sigma^2).$$

First, we list the steps of the proof and then we present a detailed development for each of them:

1. Step 1: Express $\frac{\partial \bar{R}_{2,\infty}(\mathbf{P}_2,\sigma_2^2)}{\partial \sigma_2^2}$ as function of the Stieltjes transforms $G_{M_1}(z)$ and $G_M(z)$,
2. Step 2: Obtain $G_{M_1}(z)$,
3. Step 3: Obtain $G_M(z)$,
4. Step 4: Integrate $\frac{\partial \bar{R}_{2,\infty}(\mathbf{P}_2,\sigma_2^2)}{\partial \sigma_2^2}$ to obtain $\bar{R}_{2,\infty}(\mathbf{P}_2,\sigma_2^2)$.

Step 1: Express $\frac{\partial \bar{R}_{2,\infty}(\mathbf{P}_2,\sigma_2^2)}{\partial \sigma_2^2}$ as a function of the Stieltjes transforms $G_{M_1}(z)$ and $G_M(z)$.
Using (2.18) and (2.17), the opportunistic rate per receive antenna $\bar{R}_2$ can be re-written as follows

$$\begin{aligned}\bar{R}_2(\mathbf{P}_2,\sigma_2^2) &= \tfrac{1}{N_2}\log_2\left|\mathbf{I}_{N_2}+\mathbf{Q}^{-\frac{1}{2}}\mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H\mathbf{Q}^{-\frac{1}{2}}\right| \\ &= \tfrac{1}{N_2}\log_2\left|\sigma_2^2\mathbf{I}_{N_2}+\mathbf{M}_1+\mathbf{M}_2\right|-\tfrac{1}{N_2}\log_2\left|\sigma_2^2\mathbf{I}_{N_2}+\mathbf{M}_1\right|,\end{aligned} \quad \text{(C.0.1)}$$

with $\mathbf{M}_1 \triangleq \mathbf{H}_{21}\mathbf{V}_{H_{11}}\mathbf{P}_1\mathbf{V}_{H_{11}}^H\mathbf{H}_{21}^H$, $\mathbf{M}_2 \triangleq \mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H$, and $\mathbf{M} = \mathbf{M}_1 + \mathbf{M}_2$. Matrices $\mathbf{M}$ and $\mathbf{M}_1$ are Hermitian Gramian matrices with eigenvalue decomposition $\mathbf{M} = \mathbf{U}_M\mathbf{\Lambda}_M\mathbf{U}_M^H$ and $\mathbf{M}_1 = \mathbf{U}_{M_1}\mathbf{\Lambda}_{M_1}\mathbf{U}_{M_1}^H$, respectively. Matrix $\mathbf{U}_M$ and $\mathbf{U}_{M_1}$ are $N_2\times N_2$ unitary matrices, and $\mathbf{\Lambda}_M = \mathrm{diag}(\lambda_{M,1},\ldots,\lambda_{M,N_2})$ and $\mathbf{\Lambda}_{M_1} = \mathrm{diag}(\lambda_{M_1,1},\ldots,\lambda_{M_1,N_2})$ are square diagonal matrices containing the eigenvalues of the matrices $\mathbf{M}$ and $\mathbf{M}_1$ in decreasing order. Expression (C.0.1) can be written as

$$\begin{aligned}\bar{R}_2(\mathbf{P}_2,\sigma_2^2) &= \frac{1}{N_2}\sum_{i=1}^{N_2}\log_2\left(\sigma_2^2+\lambda_{M,i}\right)-\log_2\left(\sigma_2^2+\lambda_{M_1,i}\right) \\ &= \int\log_2\left(\lambda+\sigma_2^2\right)\mathrm{d}F_M^{(N_2)}(\lambda)-\log_2\left(\lambda+\sigma_2^2\right)\mathrm{d}F_{M_1}^{(N_2)}(\lambda) \\ &\xrightarrow{a.s} \int\log_2\left(\lambda+\sigma_2^2\right)\mathrm{d}F_M(\lambda)-\int\log_2\left(\lambda+\sigma_2^2\right)\mathrm{d}F_{M_1}(\lambda),\end{aligned} \tag{C.0.2}$$

where $F_M^{(N_2)}$ and $F_{M_1}^{(N_2)}$ are respectively the empirical eigenvalue distributions of matrices $\mathbf{M}$ and $\mathbf{M}_1$ of size N_2, that converge almost surely to the asymptotic eigenvalue distributions F_M and F_{M_1}, respectively. F_M and F_{M_1} have a compact support. Indeed the empirical eigenvalue distribution of Wishart matrices $\mathbf{H}_{ij}\mathbf{H}_{ij}^H$ converges almost surely to the compactly supported Marčenko-Pastur law, and by assumption, matrices $\mathbf{V}_i\mathbf{P}_i\mathbf{V}_i^H$, $i\in\{1,2\}$ have a limit eigenvalue distribution with a compact support. Then by *Lemma 5* in [29], the asymptotic eigenvalue distribution of $\mathbf{M}_1$ and $\mathbf{M}_2$ have a compact support. The logarithm function being continuous, it is bounded on the compact supports of the asymptotic eigenvalue distributions of $\mathbf{M}_1$ and $\mathbf{M}$, therefore, the almost sure convergence in (C.0.2) could be obtained by using the bounded convergence theorem [4].
From (C.0.2), the derivative of the asymptotic rate $\bar{R}_{2,\infty}(\mathbf{P}_2,\sigma^2)$ with respect to the noise power σ_2^2 can be written as

$$\begin{aligned}\frac{\partial}{\partial\sigma_2^2}\bar{R}_{2,\infty}(\mathbf{P}_2,\sigma_2^2) &= \frac{1}{\ln 2}\left(\int\frac{1}{\sigma_2^2+\lambda}\mathrm{d}F_M(\lambda)-\int\frac{1}{\sigma_2^2+\lambda}\mathrm{d}F_{M_1}(\lambda)\right) \\ &= \frac{1}{\ln 2}\left(G_M\left(-\sigma_2^2\right)-G_{M_1}\left(-\sigma_2^2\right)\right),\end{aligned} \tag{C.0.3}$$

where $G_M(z)$ and $G_{M_1}(z)$ are the Stieltjes transforms of the asymptotic eigenvalue distributions F_M and F_{M_1}, respectively.
Step 2: Obtain $G_{M_1}(z)$
Matrix $\mathbf{M}_1$ can be written as

$$\mathbf{M}_1=\sqrt{\alpha_{21}}\mathbf{H}_{21}\mathbf{V}_{H_{11}}\frac{\mathbf{P}_1}{\alpha_{21}}\mathbf{V}_{H_{11}}^H\mathbf{H}_{21}^H\sqrt{\alpha_{21}}. \tag{C.0.4}$$

The entries of the $N_2\times M_1$ matrix $\sqrt{\alpha_{21}}\mathbf{H}_{21}$ are zero-mean i.i.d. complex Gaussian with variance $\frac{\alpha_{21}}{M_1}=\frac{1}{N_2}$, thus $\sqrt{\alpha_{21}}\mathbf{H}_{21}$ is bi-unitarily invariant. Matrix $\mathbf{V}_{H_{11}}$ is unitary, consequently $\sqrt{\alpha_{21}}\mathbf{H}_{21}\mathbf{V}_{H_{11}}$ has the same distribution as $\sqrt{\alpha_{21}}\mathbf{H}_{21}$, in particular its entries are i.i.d. with mean zero and variance $\frac{1}{N_2}$. From (2.4), $\frac{\mathbf{P}_1}{\alpha_{21}}$ is diagonal, and by assumption it has a limit eigenvalue distribution $F_{\frac{\mathbf{P}_1}{\alpha_{21}}}$. Thus we can apply *Theorem 1.1* in [104] to $\mathbf{M}_1$, in the particular case where $\mathbf{A}=\mathbf{0}_{N_2}$ to obtain the Stieltjes transform of the asymptotic eigenvalue distribution of matrix

$\mathbf{M}_1$

$$\begin{aligned} G_{M_1}(z) &= G_{\mathbf{0}_{N_2}}\Big(z-\alpha_{21}\int \tfrac{\lambda}{1+\lambda G_{M_1}(z)}\mathrm{d}F_{\frac{P_1}{\alpha_{21}}}(\lambda)\Big) \\ &= G_{\mathbf{0}_{N_2}}\Big(z-\alpha_{21}\int_{-\infty}^{\infty} \tfrac{\lambda}{1+\lambda G_{M_1}(z)}\alpha_{21} f_{P_1}(\alpha_{21}\lambda)\mathrm{d}\lambda\Big) \\ &= G_{\mathbf{0}_{N_2}}\Big(z-\int_{-\infty}^{\infty} \tfrac{t}{1+\frac{t}{\alpha_{21}}G_{M_1}(z)} f_{P_1}(t)\mathrm{d}t\Big) \\ &= G_{\mathbf{0}_{N_2}}\big(\, z-g(G_{M_1}(z))\,\big), \end{aligned} \tag{C.0.5}$$

where the function $g(u)$ is defined by

$$g(u) \triangleq \int_{-\infty}^{\infty} \tfrac{t}{1+\frac{t}{\alpha_{21}}u} f_{P_1}(t)\mathrm{d}t = \mathbb{E}\left[\tfrac{t}{1+\frac{1}{\alpha_{21}}tu}\right],$$

where the random variable t follows the c.d.f. F_{P_1}.
The square null matrix $\mathbf{0}$ has an asymptotic eigenvalue distribution $F_{\mathbf{0}}(\lambda) = \mu(\lambda)$. Thus, its Stieltjes transform is

$$G_{\mathbf{0}}(z) = \int_{-\infty}^{\infty} \frac{1}{\lambda - z}\delta(\lambda)\mathrm{d}\lambda = -\frac{1}{z}. \tag{C.0.6}$$

Then, using expressions (C.0.5) and (C.0.6), we obtain

$$G_{M_1}(z) = \frac{-1}{z - g(G_{M_1}(z))}. \tag{C.0.7}$$

Expression (C.0.7) is a fixed-point equation with unique solution when $z \in \mathbb{R}^-$ [104].
Step 3: Obtain $G_M(z)$ Recall that

$$\mathbf{M} \triangleq \mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H + \mathbf{H}_{21}\mathbf{V}_{H_{11}}\mathbf{P}_1\mathbf{V}_{H_{11}}^H\mathbf{H}_{21}^H \tag{C.0.8}$$

To obtain the Stieltjes transform G_M, we apply *Theorem 1.1* in [104] as in Step 2:

$$G_M\,(z) = G_{M_2}\,(\,z - g(G_M(z))\,)\,. \tag{C.0.9}$$

To obtain the Stieltjes transform G_{M_2} of the asymptotic eigenvalue distribution function of the matrix $\mathbf{M}_2 = \mathbf{H}_{22}\mathbf{V}_2\mathbf{P}_2\mathbf{V}_2^H\mathbf{H}_{22}^H$, we first express its S-transform as

$$\begin{aligned} S_{M_2}(z) &= S_{H_{22}V_2P_2V_2^HH_{22}^H}(z) \\ &= S_{\sqrt{\alpha_{22}}H_{22}V_2\frac{P_2}{\alpha_{22}}V_2^HH_{22}^H\sqrt{\alpha_{22}}}(z), \end{aligned}$$

and by *Lemma 1 in [29]*:

$$S_{M_2}(z) = \left(\tfrac{z+1}{z+\alpha_{22}}\right) S_{\sqrt{\alpha_{22}}H_{22}^HH_{22}\sqrt{\alpha_{22}}V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right),$$

and by *Theorem 1 in [65]*:

$$\begin{aligned} S_{M_2}(z) &= \left(\tfrac{z+1}{z+\alpha_{22}}\right) S_{\sqrt{\alpha_{22}}H_{22}^HH_{22}\sqrt{\alpha_{22}}}\left(\tfrac{z}{\alpha_{22}}\right) S_{V_2\frac{\mathbf{P}_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right) \\ &= \left(\tfrac{z+1}{z+\alpha_{22}}\right)\left(\tfrac{1}{1+\alpha_{22}\frac{z}{\alpha_{22}}}\right) S_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right) \\ &= \left(\tfrac{1}{z+\alpha_{22}}\right) S_{V_2\frac{\mathbf{P}_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right). \end{aligned} \tag{C.0.10}$$

The S-transforms $S_{M_2}(z)$ and $S_{V_2\mathbf{P}_2V_2^H}\left(\frac{z}{\alpha}\right)$ in expression (C.0.10) can be written as functions of their Υ-transforms:

$$S_{M_2}(z) = \tfrac{1+z}{z}\Upsilon^{-1}_{M_2}(z), \text{from (A.0.3)} \tag{C.0.11}$$

$$\begin{aligned} S_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right) &= \tfrac{1+\frac{z}{\alpha_{22}}}{\frac{z}{\alpha_{22}}}\Upsilon^{-1}_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right), \text{from (A.0.3)} \\ &= \tfrac{\alpha_{22}+z}{z}\Upsilon^{-1}_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\tfrac{z}{\alpha_{22}}\right). \end{aligned} \tag{C.0.12}$$

Then, plugging (C.0.11) and (C.0.12) into (C.0.10) yields

$$\Upsilon^{-1}_{M_2}(z) = \left(\frac{1}{1+z}\right)\Upsilon^{-1}_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(\frac{z}{\alpha_{22}}\right). \tag{C.0.13}$$

Now, using the relation (A.0.5) between the Υ-transform and the Stieltjes transform, we write

$$G_{M_2}(z) = \left(\frac{-1}{z}\right)\left(\Upsilon_{M_2}\left(\frac{1}{z}\right)+1\right), \tag{C.0.14}$$

and from (C.0.9), we obtain

$$G_M(z) = \left(\tfrac{-1}{z-g(G_M(z))}\right)\left(\Upsilon_{M_2}\left(\tfrac{1}{z-g(G_M(z))}\right)+1\right). \tag{C.0.15}$$

We handle (C.0.15) to obtain $G_M(z)$ as a function of $\Upsilon_{V_2P_2V_2^H}(z)$:

$$\begin{aligned} \Upsilon_{M_2}\left(\tfrac{1}{z-g(G_M(z))}\right) &= -1-(\,z-g(G_M(z))\,)\,G_M(z) \\ \tfrac{1}{z-g(G_M(z))} &= \Upsilon^{-1}_{M_2}(-1-(\,z-g(G_M(z))\,)\,G_M(z)) \\ \tfrac{1}{z-g(G_M(z))} &= \tfrac{-1}{(\,z-g(G_M(z))\,)\,G_M(z)} \\ &\quad \Upsilon^{-1}_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(-\tfrac{1+(\,z-g(G_M(z))\,)\,G_M(z)}{\alpha_{22}}\right) \\ -G_M(z) &= \Upsilon^{-1}_{V_2\frac{P_2}{\alpha_{22}}V_2^H}\left(-\tfrac{1+(\,z-g(G_M(z))\,)\,G_M(z)}{\alpha_{22}}\right) \\ \Upsilon_{V_2\frac{P_2}{\alpha_{22}}V_2^H}(-G_M(z)) &= -\tfrac{1+(\,z-g(G_M(z))\,)\,G_M(z)}{\alpha_{22}} \\ G_M(z) &= \left(-\tfrac{1}{z-g(G_M(z))}\right) \\ &\quad \left(1+\alpha_{22}\Upsilon_{V_2\frac{P_2}{\alpha_{22}}V_2^H}(-G_M(z))\right). \end{aligned} \tag{C.0.16}$$

From the definition of the Υ-transform (A.0.2), it follows that

$$\begin{aligned} \alpha_{22}\Upsilon_{V_2\frac{P_2}{\alpha_{22}}V_2^H}(-G_M(z)) &= \alpha_{22}\int \tfrac{-G_M(z)\lambda}{1+G_M(z)\lambda}\mathrm{d}F_{V_2\frac{P_2}{\alpha_{22}}V_2^H}(\lambda) \\ &= \int \tfrac{-\alpha_{22}G_M(z)\lambda}{1+G_M(z)\lambda}\alpha_{22}f_{V_2P_2V_2^H}(\alpha_{22}\lambda)\mathrm{d}\lambda \\ &= \int \tfrac{-G_M(z)t}{1+G_M(z)\frac{t}{\alpha_{22}}}f_{V_2P_2V_2^H}(t)\mathrm{d}t. \end{aligned} \tag{C.0.17}$$

Using (C.0.17) in (C.0.16), we have

$$G_M(z) = \left(-\tfrac{1}{z-g(G_M(z))}\right)(1-G_M(z)\,h(G_M(z))\,), \tag{C.0.18}$$

with the function $h(u)$ defined as follows

$$h(u) \triangleq \int \frac{t}{1+\frac{u}{\alpha_{22}}t}\mathrm{d}F_{V_2P_2V_2^H}(t) = \mathbb{E}\left[\frac{p_2}{1+\frac{1}{\alpha_{22}}p_2u}\right],$$

where the random variable p_2 follows the distribution $F_{V_2 P_2 V_2^H}$.
Factorizing $G_M(z)$ in (C.0.18) finally yields

$$G_M(z) = \frac{-1}{z - g(G_M(z)) - h(G_M(z))}. \tag{C.0.19}$$

Expression (C.0.19) is a fixed point equation with unique solution when $z \in \mathbb{R}_-$ [104].
Step 4: Integrate $\frac{\partial \bar{R}_2(\mathbf{P}_2, \sigma_2^2)}{\partial \sigma_2^2}$ to obtain $\bar{R}_2(\mathbf{P}_2, \sigma_2^2)$ in the RLNA.
From (C.0.3), we have that

$$\frac{\partial}{\partial \sigma_2^2} \bar{R}_{2,\infty}(\mathbf{P}_2, \sigma_2^2) = \frac{1}{\ln 2}\left(G_M\left(-\sigma_2^2\right) - G_{M_1}\left(-\sigma_2^2\right)\right). \tag{C.0.20}$$

Moreover, it is know that if $\sigma_2^2 \to \infty$ no reliable communication is possible and thus, $\bar{R}_{2,\infty} = 0$. Hence, the asymptotic rate of the opportunistic link can be obtained by integrating expression (C.0.20)

$$\bar{R}_{2,\infty} = \frac{-1}{\ln 2} \int_{\sigma_2^2}^{\infty} \left(G_M\left(-z\right) - G_{M_1}\left(-z\right)\right) \mathrm{d}z, \tag{C.0.21}$$

which ends the proof.

Appendix D

Proof of Prop. 3.5.1

In this appendix we provide a proof of Prop. 3.5.1. The proof is separated in two steps. First, we show that a power allocation vector $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2) \in \mathcal{P}$ of the form

$$\mathbf{p}_1 = (p_{11}, p_{\max} - p_{11}) \text{ and } \mathbf{p}_2 = (p_{\max} - p_{22}, p_{22}),$$

where, $p_{11} \in]0, p_{\max}[$ and $p_{22} \in]0, p_{\max}[$ is never observed at the NE of the game $\mathcal{G}_{(a)}$. Second, we show that if a given PA vector $\mathbf{p}^{\dagger}$ is an NE, then depending on the channel gain vector $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22})$, $\mathbf{p}^{\dagger} \in \mathcal{P}_n$, where each set $\mathcal{P}_n$, $n \in \{1, \ldots, 4\}$, is described hereunder. **Proof**: **First Step:** Assume that the action profile $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2)$, with $\mathbf{p}_1 = (p_{11}, p_{12})$ and $\mathbf{p}_2 = (p_{21}, p_{22})$ is an NE of the game $\mathcal{G}_{(a)}$, and assume that for all $(k, s) \in \mathcal{K} \times \mathcal{S}$, $p_{k,s} > 0$, with strict inequality. Then, from (3.14), it can be implied that $\forall (k, s) \in \mathcal{K} \times \mathcal{S}$,

$$\frac{1}{\beta_k} > \frac{\sigma^2 + p_{-k,s} g_{-k,s}}{g_{k,s}} \quad \text{and} \quad \frac{1}{\beta_k} > \frac{\sigma^2 + p_{-k,-s} g_{-k,-s}}{g_{k,-s}}, \tag{D.0.1}$$

where β_k is the water-level of player k at the NE. Hence, the action profile $\mathbf{p}_k$, with $k \in \mathcal{K}$, can be written as follows:

$$\mathbf{p}_1 = (p_{11}, p_{\max} - p_{11}) \text{ and } \mathbf{p}_2 = (p_{\max} - p_{22}, p_{22})$$

where, $p_{11} \in]0, p_{\max}[$ and $p_{22} \in]0, p_{\max}[$. More specifically, we have that: $\forall k \in \mathcal{K}$,

$$\begin{array}{rcl} p_{k,k} & = & \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{-k,k}\left(p_{\max} - p_{-k,-k}\right)}{g_{k,k}} + \frac{\sigma^2 + g_{-k,-k} p_{-k,-k}}{g_{k,-k}}\right) \\ p_{k,-k} & = & p_{\max} - p_{k,k}. \end{array} \tag{D.0.2}$$

Using a matrix notation, the system of equations (D.0.2) can be written as follows:

$$\left(\begin{array}{cc} 2g_{11}g_{12} & -(g_{22}g_{11}+g_{21}g_{12}) \\ -(g_{22}g_{11}+g_{21}g_{12}) & 2g_{11}g_{12} \end{array}\right)\left(\begin{array}{c} p_{11} \\ p_{22} \end{array}\right) = \mathbf{A},$$

where, the matrix $\mathbf{A}$ is

$$\mathbf{A} = \left(\begin{array}{c} p_{\max} g_{12}\left(g_{11} - g_{21}\right) + \sigma^2\left(g_{11} - g_{12}\right) \\ p_{\max} g_{21}\left(g_{22} - g_{12}\right) + \sigma^2\left(g_{22} - g_{21}\right) \end{array}\right). \tag{D.0.3}$$

Assuming that the set of channels $\{g_{11}, g_{12}, g_{21}, g_{22}\}$ always satisfy the condition:

$$g_{12}g_{21} - g_{11}g_{22} \neq 0, \tag{D.0.4}$$

the system of equations (D.0.2) has the following unique solution: $\forall k \in \mathcal{K}$,

$$p_{k,k} = \frac{p_{\max} g_{-k,k}\left(g_{k,-k}+g_{-k,-k}\right)+\sigma^2\left(g_{-k,k}+g_{-k,-k}\right)}{g_{12}g_{21}-g_{11}g_{22}}. \tag{D.0.5}$$

Note that if $g_{12}g_{21}-g_{11}g_{22} < 0$, then $\forall k \in \mathcal{K}$, $p_{k,k} < 0$, and, if $g_{12}g_{21}-g_{11}g_{22} > 0$, then $\forall k \in \mathcal{K}$, $p_{k,k} > p_{\max}$, which contradicts the initial power constraints (2.3). Hence, any vector $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2)$, with $\mathbf{p}_1 = (p_{11}, p_{\max} - p_{11})$ and $\mathbf{p}_2 = (p_{\max} - p_{22}, p_{22})$, such that

$$p_{11} \in \left]0, p_{\max}\right[\quad \text{and} \quad p_{22} \in \left]0, p_{\max}\right[, \tag{D.0.6}$$

is not an NE for the game $\mathcal{G}_{(a)}$.

Second Step: If the condition (D.0.6) is never observed at the NE, any action profile $\mathbf{p}^\dagger = \left(\mathbf{p}_1^\dagger, \mathbf{p}_2^\dagger\right)$, which is an NE, satisfies that $\mathbf{p}^\dagger \in \mathcal{P}^\dagger$, where,

$$\begin{aligned} \mathcal{P}^\dagger &= \mathcal{P} \setminus \{\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathbb{R}_+^4 : \\ &\quad p_{11} \in \left]0, p_{\max}\right[\text{ and } p_{22} \in \left]0, p_{\max}\right[\} \\ &= \bigcup_{n=1}^{8} \mathcal{P}_i^\dagger, \end{aligned} \tag{D.0.7}$$

where the sets $\mathcal{P}_n^\dagger \subset \mathcal{P}$, for all $n \in \{1, \ldots, 4\}$ are described as follows

$$\begin{aligned} \mathcal{P}_1^\dagger &= \{\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathbb{R}_+^4 : p_{11} = 0, \text{ and } p_{22} \in \left]0, p_{\max}\right[\}, \\ \mathcal{P}_2^\dagger &= \{\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathbb{R}_+^4 : p_{11} = p_{\max}, \text{ and } p_{22} \in \left]0, p_{\max}\right[\}, \\ \mathcal{P}_3^\dagger &= \{\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathbb{R}_+^4 : p_{11} \in \left]0, p_{\max}\right[\text{ and } p_{22} = 0\}, \\ \mathcal{P}_4^\dagger &= \{\mathbf{p} = (p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22}) \in \mathbb{R}_+^4 : p_{11} \in \left]0, p_{\max}\right[\text{ and } p_{22} = p_{\max}\}, \end{aligned}$$

whereas, the sets $\mathcal{P}_n^\dagger \subset \mathcal{P}$, for all $n \in \{5, \ldots, 8\}$, are the singletons: $\mathcal{P}_5^\dagger = \{\mathbf{p} = (p_{\max}, 0, 0, p_{\max})\}$, $\mathcal{P}_6^\dagger = \{\mathbf{p} = (p_{\max}, 0, p_{\max}, 0)\}$, $\mathcal{P}_7^\dagger = \{\mathbf{p} = (0, p_{\max}, 0, p_{\max})\}$, $\mathcal{P}_8^\dagger = \{\mathbf{p} = (0, p_{\max}, p_{\max}, 0)\}$.

If $\mathbf{p}^\dagger$ is an NE, then it satisfies that

$$\forall k \in \mathcal{K}, \quad \mathrm{BR}_k\left(\mathbf{p}_{-k}^\dagger\right) = \mathbf{p}_k^\dagger, \tag{D.0.8}$$

where the best response function of player $k \in \mathcal{K}$, denoted by BR_k is given by (3.14). For instance, assume that $\mathbf{p}^\dagger \in \mathcal{P}_1^\dagger$, i.e., $\mathbf{p}_1^\dagger = (0, p_{\max})$ and $\mathbf{p}_2^\dagger = \left(p_{\max} - p_{22}^\dagger, p_{22}^\dagger\right)$, with $p_{22}^\dagger \in \left]0, p_{\max}\right[$. Then, from (D.0.8) with $k = 2$, we have that:

$$\frac{1}{\beta_2} > \frac{\sigma^2 + g_{12}p_{\max}}{g_{11}} \text{ and } \frac{1}{\beta_2} > \frac{\sigma^2}{g_{21}}, \tag{D.0.9}$$

which implies that

$$p_{22} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{11}} + \frac{\sigma^2}{g_{21}}\right) > 0 \tag{D.0.10}$$

and

$$p_{22} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{11}} + \frac{\sigma^2}{g_{21}}\right) < p_{\max}. \quad \text{(D.0.11)}$$

The inequality (D.0.10), holds only if: $\frac{g_{21}}{g_{22}} < \frac{1+\frac{p_{\max}g_{12}}{\sigma^2}}{1+\frac{p_{\max}g_{21}}{\sigma^2}}$, whereas inequality (D.0.11) holds only if $\frac{g_{21}}{g_{22}} > \frac{1}{1+\mathrm{SNR}(g_{12}+g_{22})}$.

Similarly, from (D.0.8) with $k = 1$, we have that:

$$\frac{1}{\beta_1} \leqslant \frac{\sigma^2 + g_{21}\left(p_{\max} - p_{22}\right)}{g_{11}} \text{ and } \frac{1}{\beta_1} \geqslant \frac{\sigma^2 + g_{22}p_{22}}{g_{12}}, \quad \text{(D.0.12)}$$

where β_1 is the water-level of player 1 at the NE and p_{22} is given by (D.0.10). Now, since β_1 saturates the power constraint (2.3), (D.0.12) implies that:

$$\frac{\sigma^2 + g_{21}\left(p_{\max} - p_{22}\right)}{g_{11}} - \frac{\sigma^2 + g_{22}p_{22}}{g_{12}} \leqslant p_{\max}. \quad \text{(D.0.13)}$$

Then, replacing (D.0.10) in (D.0.13), we obtain that expressions (D.0.12) are satisfied only if:

$$\frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{22}}. \quad \text{(D.0.14)}$$

Hence, we can conclude that when the vector $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_1$, the NE is of the form $(p_{11}, p_{\max} - p_{11}, p_{\max} - p_{22}, p_{22})$, with

$$p_{11} = 0 \text{ and} \quad \text{(D.0.15)}$$

$$p_{22} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{11}} + \frac{\sigma^2}{g_{21}}\right), \quad \text{(D.0.16)}$$

as stated in Prop. 3.5.1. Now, assuming that $\mathbf{p}^\dagger \in \mathcal{P}_n^\dagger$, with $n \in \{2, 3, 4\}$, leads to the conditions of the other types of NE, i.e., the sets $\mathcal{B}_n$. To complete the proof, assume now that $\mathbf{p}^\dagger \in \mathcal{P}_5^\dagger$, i.e., $\mathbf{p}_1 = (p_{\max}, 0)$ and $\mathbf{p}_2 = (0, p_{\max})$. Then, from (D.0.8) with $k = 1$, we have that:

$$\frac{\sigma^2 + g_{22}p_{\max}}{g_{12}} \leqslant \frac{1}{\beta_1} \leqslant \frac{\sigma^2}{g_{11}}, \quad \text{(D.0.17)}$$

whereas with $k = 2$,

$$\frac{\sigma^2 + g_{11}p_{\max}}{g_{21}} \leqslant \frac{1}{\beta_2} \leqslant \frac{\sigma^2}{g_{22}}. \quad \text{(D.0.18)}$$

Hence, from (D.0.17) - (D.0.18) and assuming that β_k, for all $k \in \mathcal{K}$, saturates the power constraints (2.3), we have that

$$\frac{\sigma^2 + g_{22}p_{\max}}{g_{12}} - \frac{\sigma^2}{g_{11}} > p_{\max}, \quad \text{(D.0.19)}$$

$$\frac{\sigma^2 + g_{11}p_{\max}}{g_{21}} - \frac{\sigma^2}{g_{22}} > p_{\max}. \quad \text{(D.0.20)}$$

The inequality (D.0.19) holds only if $\frac{g_{11}}{g_{12}} \geqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}$ and inequality (D.0.20) holds only if $\frac{g_{21}}{g_{22}} \geqslant \frac{1+\mathrm{SNR}g_{11}}{1+\mathrm{SNR}g_{22}}$. Then, finally, when the vector $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B}_5$, the NE is of the form $\mathbf{p}_1 = (p_{\max}, 0)$ and $\mathbf{p}_2 = (0, p_{\max})$. Assuming that $\mathbf{p}^\dagger \in \mathcal{P}_n$, with $n \in \{6, 7, 8\}$, and carrying out a similar analysis as presented above for $n = 4$ will lead us to the conditions to observe the other types of NE, which ends the proof. □

Appendix E

Proof of Prop. 3.5.2

Proof: By solving the system of equations (3.30) according to the conditions in Prop 3.5.2, one can immediately imply that none of the players transmits at maximum power over a single channel. On the contrary, each player will always allocate a non-zero power to both of its channels. Hence, one can write the best response for player $k = 1$ as follows,

$$p_{11} = \frac{1}{\beta_1} - \frac{\sigma^2}{g_{11}} - \frac{1}{\alpha} p_{2,1} > 0 \tag{E.0.1}$$

$$p_{12} = \frac{1}{\beta_1} - \frac{\sigma^2}{g_{12}} - \frac{1}{\alpha} p_{2,2} > 0. \tag{E.0.2}$$

Now, using the fact that for all $k \in \mathcal{K}$, $p_{k,1} + p_{k,2} = p_{\max}$, we have that:

$$\frac{1}{\beta_1} = \frac{1}{2}\left(p_{\max}\left(1 + \frac{1}{\alpha}\right) + \sigma^2\left(\frac{1}{g_{11}} + \frac{1}{g_{12}}\right)\right), \tag{E.0.3}$$

and thus, from (E.0.1) and (E.0.2) it yields,

$$\begin{cases} p_{11} = \frac{1}{2}\left(p_{\max}\left(1 - \frac{1}{\alpha}\right) + \sigma^2\left(\frac{1}{g_{12}} - \frac{1}{g_{11}}\right)\right) + \frac{1}{\alpha} p_{22} \\ p_{12} = p_{\max} - p_{11}. \end{cases} \tag{E.0.4}$$

One can verify that solving the system of equations (3.30) subject to the conditions in Prop. 3.5.2 and following the same reasoning as above for player $k = 2$ leads us to the following equalities:

$$\begin{cases} p_{22} = \frac{1}{2}\left(p_{\max}(1 + \alpha) + \sigma^2\left(\frac{1}{g_{21}} - \frac{1}{g_{22}}\right)\right) + \alpha p_{11}, \\ p_{21} = p_{\max} - p_{22}, \end{cases} \tag{E.0.5}$$

where the first equations in both (E.0.4) and (E.0.5) are identical, which implies that any two PA vectors $\mathbf{p}_1 = (p_{11}, p_{\max} - p_{11}) \in \mathcal{P}_1^{(a)}$ and $\mathbf{p}_2 = (p_{\max} - p_{22}, p_{22}) \in \mathcal{P}_2^{(a)}$ satisfying the conditions in Prop. 3.5.2 are one NE of the game $\mathcal{G}_{(a)}$. $\square$

Appendix F

Proof of Prop. 3.5.5

Proof: The proof follows from the fact that in the low SNR regime, i.e., SNR $\rightarrow 0$, one can write that:

$$\lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_1 = \lim_{\mathrm{SNR}\rightarrow 0} \mathcal{B}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 \text{ and } \frac{g_{21}}{g_{22}} \leqslant 1\} \quad \text{(F.0.1)}$$

$$\lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_2 = \lim_{\mathrm{SNR}\rightarrow 0} \mathcal{B}_2 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant 1 \text{ and } \frac{g_{21}}{g_{22}} \geqslant 1\} \quad \text{(F.0.2)}$$

$$\lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_3 = \lim_{\mathrm{SNR}\rightarrow 0} \mathcal{B}_3 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant 1 \text{ and } \frac{g_{21}}{g_{22}} \leqslant 1\} \quad \text{(F.0.3)}$$

$$\lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_4 = \lim_{\mathrm{SNR}\rightarrow 0} \mathcal{B}_4 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant 1 \text{ and } \frac{g_{21}}{g_{22}} \geqslant 1\} \quad \text{(F.0.4)}$$

and moreover,

$$\forall n \in \{5, \ldots, 8\}, \quad \lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_n = \emptyset, \quad \text{(F.0.5)}$$

Thus, from (F.0.1) - (F.0.5), Prop. 3.5.1 and Prop. 3.5.3, one can immediately conclude that the PA vector given by (3.39) and (3.40) is the NE of both games $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$. The uniqueness of the NE in the game $\mathcal{G}_{(a)}$ holds with probability one, independently of SNR level (Prop. 3.5.1). Conversely, in the game $\mathcal{G}_{(b)}$, the NE is unique in the low SNR regime, due to the fact that

$$\lim_{\mathrm{SNR}\rightarrow 0} \mathcal{A}_5 \quad = \quad \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} = 1 \text{ and } \frac{g_{21}}{g_{22}} = 1\},$$

and since for all $(k, s) \in \mathcal{K} \times \mathcal{S}$, $g_{k,s}$ is a realization of a random variable drawn from a continuous probability distribution, we have that

$$\exists \mathrm{SNR}_0, \forall \mathrm{SNR} \leqslant \mathrm{SNR}_0, \quad \Pr\left(\mathbf{g} \in \mathcal{A}_5\right) \leqslant \epsilon\left(\mathrm{SNR}\right), \quad \text{(F.0.6)}$$

where, $\lim_{\mathrm{SNR}\rightarrow\infty} \epsilon\left(\mathrm{SNR}\right) = 0$. Thus, since the NE of the game $\mathcal{G}_{(b)}$ is not unique only if $\mathbf{g} \in \mathcal{A}_5$, we have that in the high SNR regime, the NE in the game $\mathcal{G}_{(b)}$ is unique with probability one, which completes the proof. $\square$

Appendix G

Proof of Prop. 3.5.6

In this appendix, we provide the proof of Prop. 3.5.6, which states that at the high SNR regime there always exists an NE action profile in the game $\mathcal{G}_{(b)}$, which leads to a better global performance than the unique NE of the game $\mathcal{G}_{(a)}$.

Before a formal proof, we introduce two lemmas which are used in the proof.

Lemma G.0.2 *In the high SNR regime, the game $\mathcal{G}_{(a)}$ possesses a unique NE, which can be of six different types depending on the channel realizations $\{g_{i,j}\}_{\forall (i,j)\in\mathcal{K}\times\mathcal{P}}$:*

- *Equilibrium 1: if $\mathbf{g} \in \mathcal{B'}_1 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \geqslant \frac{g_{11}}{g_{22}}, \text{ and } \frac{g_{21}}{g_{22}} \leqslant \frac{g_{11}}{g_{22}}\}$, then, $p_{11}^{\dagger} = p_{\max}$ and $p_{22}^{\dagger} = p_{\max}$.*
- *Equilibrium 4: if $\mathbf{g} \in \mathcal{B'}_4\{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} \leqslant \frac{g_{21}}{g_{12}}, \text{ and } \frac{g_{21}}{g_{22}} \geqslant \frac{g_{21}}{g_{12}}\}$, then, $p_{11}^{\dagger} = 0$ and $p_{22}^{\dagger} = 0$.*
- *Equilibrium 5: if $\mathbf{g} \in \mathcal{B'}_5 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} > \frac{g_{21}}{g_{22}}, \text{ and } \frac{g_{21}}{g_{22}} > \frac{g_{11}}{g_{22}}\}$, then, $p_{11}^{\dagger} = p_{\max}$ and $p_{22}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{22}} + \frac{\sigma^2 + g_{11}p_{\max}}{g_{21}}\right)$.*
- *Equilibrium 6: if $\mathbf{g} \in \mathcal{B'}_6 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} > \frac{g_{21}}{g_{22}}, \text{ and } \frac{g_{11}}{g_{12}} < \frac{g_{11}}{g_{22}}\}$, then, $p_{11}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2}{g_{11}} + \frac{\sigma^2 + p_{\max}g_{22}}{g_{12}}\right)$ and $p_{22}^{\dagger} = p_{\max}$.*
- *Equilibrium 7: if $\mathbf{g} \in \mathcal{B'}_7 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} < \frac{g_{21}}{g_{22}}, \text{ and } \frac{g_{11}}{g_{12}} > \frac{g_{21}}{g_{12}}\}$, then, $p_{11}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + p_{\max}g_{21}}{g_{11}} + \frac{\sigma^2}{g_{12}}\right)$ and $p_{22}^{\dagger} = 0$.*
- *Equilibrium 8: if $\mathbf{g} \in \mathcal{B'}_8 = \{\mathbf{g} \in \mathbb{R}_+^4 : \frac{g_{11}}{g_{12}} < \frac{g_{21}}{g_{22}}, \text{ and } \frac{g_{21}}{g_{22}} < \frac{g_{21}}{g_{12}}\}$, then, $p_{11}^{\dagger} = 0$ and $p_{22}^{\dagger} = \frac{1}{2}\left(p_{\max} - \frac{\sigma^2 + g_{12}p_{\max}}{g_{22}} + \frac{\sigma^2}{g_{21}}\right)$.*

The proof of Prop. G.0.2 follows the same reasoning of the proof of Prop 3.5.1 assuming that SNR $\to \infty$.

In Lemma G.0.2, we have deliberately indexed the different types of NE using the set $\{1, 4, \ldots, 8\}$, for the ease of notation. In the following lemma, we describe the set of NE of the game $\mathcal{G}_{(b)}$ in the high SNR regime.

Lemma G.0.3 *In the high SNR regime, the game $\mathcal{G}_{(b)}$ always possesses two NE action profiles:*

$$\mathbf{p}_1^{*,1} = (0, p_{\max}) \text{ and } \mathbf{p}_2^{*,1} = (p_{\max}, 0) \tag{G.0.1}$$

and

$$\mathbf{p}_1^{*,4} = (p_{\max}, 0) \text{ and } \mathbf{p}_2^{*,4} = (0, p_{\max}), \tag{G.0.2}$$

independently of the channel realizations.

Proof: In the high SNR, i.e., SNR $\to \infty$, one can write

$$\lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_1 = \{\mathbf{g}\in\mathbb{R}_+^4 : \tfrac{g_{11}}{g_{12}} \geqslant 0 \text{ and } \tfrac{g_{21}}{g_{22}} \leqslant +\infty\} \tag{G.0.3}$$

$$\lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_2 = \{\mathbf{g}\in\mathbb{R}_+^4 : \tfrac{g_{11}}{g_{12}} \geqslant +\infty \text{ and } \tfrac{g_{21}}{g_{22}} \geqslant +\infty\} \tag{G.0.4}$$

$$\lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_3 = \{\mathbf{g}\in\mathbb{R}_+^4 : \tfrac{g_{11}}{g_{12}} \leqslant 0 \text{ and } \tfrac{g_{21}}{g_{22}} \leqslant 0\} \tag{G.0.5}$$

$$\lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_4 = \{\mathbf{g}\in\mathbb{R}_+^4 : \tfrac{g_{11}}{g_{12}} \leqslant +\infty \text{ and } \tfrac{g_{21}}{g_{22}} \geqslant 0\}. \tag{G.0.6}$$

Note that from (G.0.4) and (G.0.5), one immediately imply that

$$\Pr\left(\mathbf{g} \in \lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_2\right) = \Pr\left(\mathbf{g} \in \lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_3\right) = 0,$$

since channel gains are random variables drawn from continuous probability distributions. Conversely,

$$\Pr\left(\mathbf{g} \in \lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_1\right) = \Pr\left(\mathbf{g} \in \lim_{\mathrm{SNR}\to+\infty} \mathcal{A}_4\right) = 1.$$

Hence, from Prop. 3.5.3, we imply that both $\mathbf{p}^{(*,1)}$ and $\mathbf{p}^{(*,4)}$ are NE action profiles of the game $\mathcal{G}_{(b)}$ in the high SNR regime regardless of the exact channel realizations $\{g_{i,j}\}_{\forall (i,j)\in\mathcal{K}\times\mathcal{P}}$, which completes the proof. □

From Lemma G.0.2 and Lemma G.0.3, it is easy to see that if $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B'}_n$, with $n \in \{1, 4\}$, then (3.41) holds since $\mathbf{p}^{\dagger}$ and at least one of the NE action profiles $\mathbf{p}^{*,n}$, with $n \in \{1, 4\}$ are identical. In the cases where $\mathbf{g} = (g_{11}, g_{12}, g_{21}, g_{22}) \in \mathcal{B'}_n$, with $n \in \{5, \ldots, 8\}$, we prove by inspection that in all the cases condition (3.41) always holds for both NE action profiles $\mathbf{p}^{(*,1)}$ and $\mathbf{p}^{(*,4)}$. For instance, assume that $\mathbf{g} \in \mathcal{B'}_5$. Then, we have that the unique NE of the game $\mathcal{G}_{(b)}$ is $\mathbf{p}^{\dagger} = (p_{\max} - p_{22}, p_{22})$, with $p_{22} = \frac{1}{2}\left(p_{\max} + \frac{\sigma^2 + p_{\max} g_{11}}{g_{21}} - \frac{\sigma^2}{g_{22}}\right)$ (Prop. 3.5.1). Define the function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ a s follows: $\psi(x) = 1 + \mathrm{SNR}x$ and denote by $\Delta_1(\mathrm{SNR})$, the difference between the sum transmission rates achieved by playing $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$, with respect to the NE $\mathbf{p}^{*,1}$ at SNR level SNR, i.e.,

$$\begin{aligned} \Delta_1(\mathrm{SNR}) &= u_1\left(\mathbf{p}^{*,1}\right) + u_2\left(\mathbf{p}^{*,1}\right) - \left(u_1\left(\mathbf{p}^{\dagger}\right) + u_2\left(\mathbf{p}^{\dagger}\right)\right) \\ &= 2\log_2(2) - 2\log_2\left(1 + \frac{g_{21}}{g_{22}}\frac{\psi(g_{22})}{\psi(g_{11})}\right) - \log_2\left(1 + \frac{g_{22}}{g_{21}}\frac{\psi(g_{11})}{\psi(g_{22})}\right) \\ &\quad + \log_2\left(\frac{g_{21}}{g_{22}} + \psi(g_{21} - g_{11})\right). \end{aligned} \tag{G.0.7}$$

Note that if $\mathbf{g} \in \mathcal{B}'_5$, then $g_{21} > g_{11}$. Hence,

$$\lim_{\mathrm{SNR}\to\infty} \Delta_1(\mathrm{SNR}) = 2\log_2(2) - 2\log_2\left(1+\frac{g_{21}}{g_{11}}\right) - \log_2\left(1+\frac{g_{11}}{g_{21}}\right) + \log_2\left(1+\frac{g_{21}}{g_{22}}+\infty\right) = \infty,$$

which justifies (3.41). Similarly, denote $\Delta_4(\mathrm{SNR})$, the difference between the sum transmission rates achieved by playing $\mathcal{G}_{(a)}$ and $\mathcal{G}_{(b)}$, with respect to the NE $\mathbf{p}^{*,4}$, i.e.,

$$\begin{aligned}\Delta_4(\mathrm{SNR}) &= u_1(\mathbf{p}^{*,4}) + u_2(\mathbf{p}^{*,4}) - \left(u_1(\mathbf{p}^{\dagger}) + u_2(\mathbf{p}^{\dagger})\right) \\ &= 2\log_2(2) - 2\log_2\left(1 + \frac{g_{22}}{g_{21}}\frac{\psi(g_{21})}{\psi(g_{12})}\right) - \log_2\left(1 + \frac{g_{21}}{g_{22}}\frac{\psi(g_{12})}{\psi(g_{21})}\right) \\ &\quad + \log_2\left(\frac{g_{22}}{g_{21}} + \psi(g_{22}-g_{12})\right). \end{aligned} \tag{G.0.8}$$

Note that if $\mathbf{g} \in \mathcal{B}'_5$, then $g_{22} > g_{12}$. Hence,

$$\lim_{\mathrm{SNR}\to\infty} \Delta_4(\mathrm{SNR}) = 2\log_2(2) - 2\log_2\left(1+\frac{g_{22}}{g_{12}}\right) - \log_2\left(1+\frac{g_{12}}{g_{22}}\right) + \log_2\left(1+\frac{g_{22}}{g_{21}}+\infty\right) = \infty,$$

which justifies (3.41). Hence, one can imply that in the high SNR regime both NE action profiles $\mathbf{p}^{*,1}$ and $\mathbf{p}^{*,2}$, satisfy (3.41) when $\mathbf{g} \in \mathcal{B}'_5$. The same result as the one obtained when $\mathbf{g} \in \mathcal{B}'_5$, is also obtained when $\mathbf{g} \in \mathcal{B}'_n$, with $n \in \{6, \ldots, 8\}$, which completes the proof.

Appendix H

Proof of Lemma 4.3.3

In this appendix, we provide the proof of Lemma 4.3.3, which states that the performance achieved by the players using a behavioral strategy $\sigma \in \Sigma$, given a particular initialization $\mathbf{a}(0) = (a_1(0), \ldots, a_K(0)) \in \mathcal{A}$ and its corresponding observation $\tilde{\mathbf{u}}(0) = (\tilde{u}_k(0), \ldots, \tilde{u}_K(0)) \in \mathbb{R}^K$, for all $k \in \mathcal{K}$, can be written as a function of its corresponding AASB, denoted by $\kappa^* \in \triangle(\mathcal{A})$.

As a first step, let the performance achieved by the player k (4.12) be written as follows,

$$\begin{aligned}
\bar{u}_k(\sigma_k, \sigma_{-k}) &= \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{(\mathbf{a}(0),\tilde{\mathbf{u}}(0),\sigma)} \left[\sum_{n=0}^{N-1} u_k(\mathbf{h}(n), a_k(n), \mathbf{a}_{-k}(n)) \right] \\
&\quad + \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \varepsilon_{k,a_k(n)}(n), \\
&= \lim_{N\to\infty} \frac{1}{N} \mathbb{E}_{(\mathbf{a}(0),\tilde{\mathbf{u}}(0),\sigma)} \left[\sum_{n=0}^{N-1} u_k(\mathbf{h}(n), a_k(n), \mathbf{a}_{-k}(n)) \right], \qquad \text{(H.0.1)} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{\mathbf{a}\in\mathcal{A}} u_k(\mathbf{h}(n), a_k, \mathbf{a}_{-k}) \prod_{j=1}^{K} \pi_{j,a_j}(n) \qquad \text{(H.0.2)} \\
&= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sum_{\mathbf{a}\in\mathcal{A}} \sum_{\mathbf{h}\in\mathcal{H}} u_k(\mathbf{h}, a_k, \mathbf{a}_{-k}) \mathbb{1}_{\{\mathbf{h}(n)=\mathbf{h}\}} \prod_{j=1}^{K} \pi_{j,a_j}(n). \qquad \text{(H.0.3)}
\end{aligned}$$

Now, let the i-th element of the set $\mathcal{H}$ be denoted by $\mathbf{h}^{(i)}$, with $i \in \{1, \ldots, H\}$. Let also the sets $\mathcal{I}_1(n), \ldots, \mathcal{I}_H(n)$ be a partition of the set $\{1, \ldots, n\} \subset \mathbb{N}$, such that, for all $i \in \{1, \ldots, H\}$ and $\forall m \in \mathcal{I}_i(n)$, it holds that $\mathbf{h}(m) = \mathbf{h}^{(i)}$. Then, it follows that for all $i \in \{1, \ldots, H\}$ and for all $\mathbf{a} = (a_1, \ldots, a_K) \in \mathcal{A}$,

$$\begin{aligned}
\frac{1}{n} \sum_{m=1}^{n} \left(\mathbb{1}_{\{\mathbf{h}(m)=\mathbf{h}^{(i)}\}} \kappa_{\mathbf{a}}(m) \right) &= \frac{|\mathcal{I}_i(n)|}{n} \left(\frac{1}{|\mathcal{I}_i(n)|} \sum_{m\in\mathcal{I}_i(n)} \kappa_{\mathbf{a}}(m) \right), \qquad \text{(H.0.4)} \\
\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \left(\mathbb{1}_{\{\mathbf{h}(m)=\mathbf{h}^{(i)}\}} \kappa_{\mathbf{a}}(m) \right) &= \left(\lim_{n\to\infty} \frac{|\mathcal{I}_i(n)|}{n} \right) \left(\lim_{n\to\infty} \frac{1}{|\mathcal{I}_i(n)|} \sum_{m\in\mathcal{I}_i(n)} \kappa_{\mathbf{a}}(m) \right),
\end{aligned}$$

where, $\kappa^* = (\kappa_{\mathbf{a}})_{\forall \mathbf{a} \in \mathcal{A}} \in \triangle(\mathcal{A})$

$$\kappa_{\mathbf{a}}(n) = \prod_{j=1}^{K} \pi_{j,a_j}(n). \tag{H.0.5}$$

Note that at each interval n, $\mathbf{h}(n)$ is a realization of a random variable following an ergodic probability distribution $(\rho_{\mathbf{h}^{(1)}}, \ldots, \rho_{\mathbf{h}^{(H)}}) \in \triangle(\mathcal{H})$. Then, by the law of large numbers (LLN), it follows that

$$\lim_{n\to\infty} \frac{|\mathcal{I}_i(n)|}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} \mathbb{1}_{\left\{\mathbf{h}(m)=\mathbf{h}^{(i)}\right\}} = \rho_{\mathbf{h}^{(i)}}. \tag{H.0.6}$$

Now, define the functions $F_{i,\mathbf{a}} : \mathbb{N} \to [0,1]$ and $F_{\mathbf{a}} : \mathbb{N} \to [0,1]$ as follows,

$$F_{i,\mathbf{a}}(m) = \frac{1}{|\mathcal{I}_i(m)|} \sum_{j\in\mathcal{I}_i(m)} \kappa_{\mathbf{a}}(j) \tag{H.0.7}$$

$$F_{\mathbf{a}}(m) = \frac{1}{m}\sum_{j=1}^{m} \kappa_{\mathbf{a}}(j). \tag{H.0.8}$$

Note, that for all $(n,m) \in \mathbb{N}^2$, it holds that

$$F_{i,\mathbf{a}}(n) - F_{i,\mathbf{a}}(n+m) = \left(\frac{|\mathcal{I}_i(n+m)| - |\mathcal{I}_i(n)|}{|\mathcal{I}_i(n)|}\right)\left(F_{i,\mathbf{a}}(n) - \frac{1}{|\mathcal{I}_i(n+m)| - |\mathcal{I}_i(n)|}\sum_{\ell\in\mathcal{I}_i(n+m)\setminus\mathcal{I}_i(n)} \kappa_{\mathbf{a}}(\ell)\right). \tag{H.0.9}$$

Thus, we can write that,

$$|F_{i,\mathbf{a}}(n) - F_{i,\mathbf{a}}(n+m)| \leqslant \left(\frac{m}{|\mathcal{I}_i(n)|}\right).$$

and thus, for all $m \in \mathbb{N}$, it holds that

$$\lim_{n\to\infty} |F_{i,\mathbf{a}}(n) - F_{i,\mathbf{a}}(n+m)| = 0. \tag{H.0.10}$$

This result implies that for all $\delta > 0$ and for all $m \in \mathbb{N}$, there always exists a number $n \in \mathbb{N}$ such that

$$|F_{i,\mathbf{a}}(n) - F_{i,\mathbf{a}}(n+m)| < \delta, \tag{H.0.11}$$

and thus, the limit $\lim_{n\to\infty} \frac{1}{|\mathcal{I}_i(n)|} \sum_{j\in\mathcal{I}_i(n)} \kappa_{\mathbf{a}}(j)$ exists. Note that the same hold for the function $F_{\mathbf{a}}$, for all $\mathbf{a} \in \mathcal{A}$. Now, note that for all $(i,\mathbf{a}) \in \{1,\ldots,K\} \times \mathcal{A}$, the terms $F_{i,\mathbf{a}}(n)$ and $F_{\mathbf{a}}(n)$, $n \in \mathbb{N}$ can be written as to stochastic approximation algorithms,

$$F_{i,\mathbf{a}}(n) = F_{i,\mathbf{a}}(n-1) + \frac{\mathbb{1}_{\left\{\mathbf{h}(n)=\mathbf{h}^{(i)}\right\}}}{|\mathcal{I}_i(m)|}\left(\kappa_{\mathbf{a}(n)} - F_{i,\mathbf{a}}(n-1)\right) \tag{H.0.12}$$

$$F_{\mathbf{a}}(n) = F_{\mathbf{a}}(n-1) + \frac{1}{t}\left(\kappa_{\mathbf{a}(n)} - F_{i,\mathbf{a}}(n-1)\right), \tag{H.0.13}$$

respectively. Then, since both dynamics can be asymptotically described by the same ordinal differential equation [14], we have that,

$$\lim_{n\to\infty} \frac{1}{|\mathcal{I}_i(n)|} \sum_{j\in\mathcal{I}_i(n)} \kappa_{\mathbf{a}}(j) \quad = \quad \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \kappa_{\mathbf{a}}(j) = \kappa^*_{\mathbf{a}}, \tag{H.0.14}$$

where $\kappa^* \in \triangle(\mathcal{A})$ is the AASB induced by the behavioral strategy σ, given the initialization $\mathbf{a}(0)$ and its respective observation $\tilde{\mathbf{u}}$. Then, it allows us to write,

$$\bar{u}_k(\sigma_k, \sigma_{-k}) \quad = \quad \sum_{\mathbf{a}\in\mathcal{A}} \sum_{\mathbf{h}\in\mathcal{H}} u_k(\mathbf{h}, a_k, \mathbf{a}_{-k}) \rho_{\mathbf{h}} \kappa^*_{\mathbf{a}},$$

which completes the proof.

Appendix I

Proof of Lemma 4.5.1

In this appendix, we provide the proof of the Lemma 4.5.1. As a first step, we rewrite equation (4.29) as follows,

$$
\begin{aligned}
\hat{u}_{k,A_k^{(n_k)}}(n) &= \frac{1}{T_{k,A_k^{(n_k)}}(n)} \sum_{s=0}^{n-2} \tilde{u}_k(s) \mathbb{1}_{\left\{a_k(s)=A_k^{(n_k)}\right\}} + \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n)} \tilde{u}_k(n-1), \\
&= \frac{1}{T_{k,A_k^{(n_k)}}(n)} \left(\frac{T_{k,A_k^{(n_k)}}(n-1)}{T_{k,A_k^{(n_k)}}(n-1)} \right) \sum_{s=0}^{n-2} \tilde{u}_k(s) \mathbb{1}_{\left\{a_k(s)=A_k^{(n_k)}\right\}} \\
&\quad + \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n)} \tilde{u}_k(n-1), \\
&= \frac{1}{T_{k,A_k^{(n_k)}}(n)} \left(\frac{T_{k,A_k^{(n_k)}}(n) - \mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n-1)} \right) \sum_{s=0}^{n-2} \tilde{u}_k(s) \mathbb{1}_{\left\{a_k(s)=A_k^{(n_k)}\right\}} \\
&\quad + \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n)} \tilde{u}_k(n-1), \\
&= \hat{u}_{k,A_k^{(n_k)}}(n-1) - \left(\frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n)} \right) \hat{u}_{k,A_k^{(n_k)}}(n-1) \\
&\quad + \frac{\mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}}}{T_{k,A_k^{(n_k)}}(n)} \tilde{u}_k(n-1), \\
&= \hat{u}_{k,A_k^{(n_k)}}(n-1) \\
&\quad + \frac{1}{T_{k,A_k^{(n_k)}}(n)} \mathbb{1}_{\left\{a_k(n-1)=A_k^{(n_k)}\right\}} \left(\tilde{u}_k(n-1) - \hat{u}_{k,A_k^{(n_k)}}(n-1) \right)
\end{aligned}
\tag{I.0.1}
$$

Now, let the function $G_k^{(1)} : \mathcal{A}_k \times \mathbb{R}^{N_k+1} \to \mathbb{R}^{N_k}$ be defined as follows for all $k \in \mathcal{K}$,

$$
G_k^{(1)}(a_k, \hat{\mathbf{u}}_k, \tilde{u}_k) = \left(\mathbb{1}_{\left\{a_k=A_k^{(1)}\right\}} \left(\tilde{u}_k - \hat{u}_{k,A_k^{(1)}} \right), \ldots, \mathbb{1}_{\left\{a_k=A_k^{(N_k)}\right\}} \left(\tilde{u}_k - \hat{u}_{k,A_k^{(N_k)}} \right) \right), \tag{I.0.2}
$$

where, $a_k \in \mathcal{A}_k$, $\hat{\mathbf{u}}_k \in \mathbb{R}^{N_k}$, and $\tilde{u}_k \in \mathbb{R}$, for all $n > 0$. Let also the function $U_k^{(1)} : \mathbb{R}^{N_k} \to \mathbb{R}^{N_k}$ be defined as follows,

$$\mathbf{U}_k^{(1)}(\hat{\mathbf{u}}_k) = \mathbb{E}_{\mathbf{h},\pi^+,\varepsilon_k}\left[G^{(1)}(a_k, \hat{\mathbf{u}}_k, u_k(\mathbf{h}, a_k, \mathbf{a}_{-k}) + \varepsilon_{k,a_k})\right]. \quad (I.0.3)$$

Denote by $U_{k,n_k}^{(1)}(\hat{\mathbf{u}}_k)$, with $n_k \in \{1, \ldots, N_k\}$, the n_k-th component of the vector $\mathbf{U}_k^{(1)}(\hat{\mathbf{u}}_k)$ such that, $\mathbf{U}_k^{(1)}(\hat{\mathbf{u}}_k) = \left(U_{k,1}^{(1)}(\hat{\mathbf{u}}_k), \ldots, U_{k,N_k}^{(1)}(\hat{\mathbf{u}}_k)\right)$ and for all $n_k \in \{1, \ldots, N_k\}$, it follows from (I.0.3) that,

$$\begin{aligned} U_{k,n_k}^{(1)}(\hat{\mathbf{u}}_k) &= \mathbb{E}_{\mathbf{h},\pi^*,\varepsilon_k}\left[\mathbb{1}_{\left\{a_k = A_k^{(n_k)}\right\}}\left(u_k\left(\mathbf{h}, A_k^{(n_k)}, \mathbf{a}_{-k}\right) + \varepsilon_{k,A_k^{(n_k)}} - \hat{u}_{k,A_k^{(n_k)}}\right)\right] \\ &= \mathbb{E}_{\pi_k^*}\left[\mathbb{1}_{\left\{a_k = A_k^{(n_k)}\right\}}\right]\left(\mathbb{E}_{\mathbf{h},\pi_{-k}^*}\left[u_k\left(\mathbf{h}, A_k^{(n_k)}, \mathbf{a}_{-k}\right)\right] + \mathbb{E}_{\varepsilon_k}\left[\varepsilon_{k,A_k^{(n_k)}}\right] - \hat{u}_{k,A_k^{(n_k)}}\right) \\ &= \pi^+_{k,A_k^{(n_k)}}\left(\bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*\right) - \hat{u}_{k,A_k^{(n_k)}}\right). \quad (I.0.4) \end{aligned}$$

Define the function $V_k^{(1)} : \mathbb{N} \setminus \{0\} \to \mathbb{R}$ as follows,

$$\mathbf{V}_k^{(1)}(n) = G_k^{(1)}(a_k(n-1), \hat{\mathbf{u}}_k(n-1), \tilde{u}_k(n-1)) - U_k^{(1)}(\hat{\mathbf{u}}_k(n-1)). \quad (I.0.5)$$

Hence, using the equations (I.0.3) and (I.0.5), we can write equation (I.0.1) as follows.

$$\hat{\mathbf{u}}_k(n) = \hat{\mathbf{u}}_k(n-1) + \frac{1}{T_{k,A_k^{(n_k)}}(n)}\left(U_k^{(1)}(\hat{\mathbf{u}}_k(n-1)) - V_k^{(1)}(n)\right). \quad (I.0.6)$$

Now, we introduce the following lemmas.

Lemma I.0.4 *For all $k \in \mathcal{K}$, the function $U_k^{(1)}$ is Lipschitz with Lipschitz constant $L = 1$.*

Proof: For all $k \in \mathcal{K}$, let $\mathbf{x}_k = (x_{k,1}, \ldots, x_{k,N_k})$ and $\mathbf{y}_k = (y_{k,1}, \ldots, y_{k,N_k})$ be two vectors in $\mathbb{R}^{N_k}$. Then,

$$\begin{aligned} \left|\left|U_k^{(1)}(\mathbf{x}) - U_k^{(1)}(\mathbf{y})\right|\right| &\leqslant L\left|\left|\mathbf{x} - \mathbf{y}\right|\right| \quad (I.0.7) \\ \left|\left|\left(\pi^+_{k,A_k^{(1)}}(x_{k,1} - y_{k,1}), \ldots, \pi^+_{k,A_k^{(N_k)}}(x_{k,N_k} - y_{k,N_k})\right)\right|\right| &\leqslant L\left|\left|\mathbf{x} - \mathbf{y}\right|\right|. \quad (I.0.8) \end{aligned}$$

Now, we recall that for all $n_k \in \{1, \ldots, N_k\}$, $0 < \pi_{k,A_k^{(n_k)}} < 1$. Hence, (I.0.8) holds for any $L \geqslant 1$. □

Lemma I.0.5 *For all $k \in \mathcal{K}$ and for all $n_k \in \{1, \ldots, N_k\}$, the term $T_{k,A_k^{(n_k)}}(n)$ defined by (4.26), satisfies that*

$$\lim_{m\to\infty} \sum_{n=0}^{m} \frac{1}{T_{k,A_k^{(n_k)}}(n)} = +\infty \quad (I.0.9)$$

$$\lim_{m\to\infty} \sum_{n=0}^{m} \left(\frac{1}{T_{k,A_k^{(n_k)}}(n)}\right)^2 < +\infty. \quad (I.0.10)$$

Proof: From assumption ($A0$), it holds that, for all $k \in \mathcal{K}$, the asymptotic probability $\pi^{+}_{k,A_k^{(n_k)}}$ of playing any action $n_k \in \{1,\ldots,N_k\}$ is non-zero. Then, all actions are played infinitely many times, and thus (I.0.9) holds $\forall k \in \mathcal{K}$ and $\forall n_k \in \{1,\ldots,N_k\}$. Now, for all $k \in \mathcal{K}$ and $\forall n_k \in \{1,\ldots,N_k\}$ let us write the empirical measure

$$\hat{\pi}_{k,A_k^{(n_k)}}(n) \;= \tfrac{1}{n}T_{k,A_k^{(n_k)}}(n) = \; \pi^{+}_{k,A_k^{(n_k)}} + \epsilon_{k,A_k^{(n_k)}}(n), \tag{I.0.11}$$

where, for all $n > 0$, the term $\epsilon_{k,A_k^{(n_k)}}(n)$ represents the estimation error at interval n. Note that from assumption ($A0$), it holds that $\lim_{n\to\infty} \epsilon_{k,A_k^{(n_k)}}(n) = 0$ and for any $n \in \mathbb{N}$, it follows that

$$-\pi^{+}_{k,A_k^{(n_k)}} < \epsilon_{k,A_k^{(n_k)}}(n) < 1 - \pi^{+}_{k,A_k^{(n_k)}}. \tag{I.0.12}$$

Thus, we can write that

$$\begin{aligned}\frac{1}{T_{k,A_k^{(n_k)}}(n)} &= \frac{1}{n}\left(\frac{1}{\pi^{+}_{k,A_k^{(n_k)}} + \epsilon_{k,A_k^{(n_k)}}(n)}\right) && \text{(I.0.13)}\\ &\leqslant \frac{1}{n}\left(\frac{1}{\pi^{+}_{k,A_k^{(n_k)}}\delta}\right) && \text{(I.0.14)}\end{aligned}$$

where, $0 < \delta < 1 + \frac{\epsilon_{k,A_k^{(n_k)}}(n)}{\pi^{+}_{k,A_k^{(n_k)}}}$ is a positive real number. Thus,

$$\begin{aligned}\lim_{N\to\infty}\sum_{n=1}^{N}\left(\frac{1}{T_{k,A_k^{(n_k)}}(n)}\right)^2 &\leqslant \lim_{N\to\infty}\sum_{n=1}^{N}\left(\frac{1}{\pi^{+}_{k,A_k^{(n_k)}}\delta}\right)^2\frac{1}{n^2} && \text{(I.0.15)}\\ &\leqslant \left(\frac{1}{\pi^{+}_{k,A_k^{(n_k)}}\delta}\right)^2\left(\lim_{N\to\infty}\sum_{n=1}^{N}\frac{1}{n^2}\right) && \text{(I.0.16)}\\ &\leqslant \left(\frac{1}{\pi^{+}_{k,A_k^{(n_k)}}\delta}\right)^2\left(\frac{\pi^2}{6}\right), && \text{(I.0.17)}\end{aligned}$$

which completes the proof. $\square$

Lemma I.0.6 *The sequence $\{\mathbf{V}_k^{(n)}\}_{n>0}$ is a martingale difference sequence with respect to the increasing family of σ-fields,*

$$\mathcal{F}^{(1)}_{k,n} \triangleq \sigma\left(\left(\hat{\mathbf{u}}_k(0), \pi(0), V_k^{(1)}(0), \mathbf{h}(0), \varepsilon(0)\right), \ldots, \left(\hat{\mathbf{u}}_k(n), \pi(n), V_k^{(1)}(n), \mathbf{h}(n), \varepsilon(n)\right)\right), \tag{I.0.18}$$

Proof: From the definition of the mapping $\mathbf{V}_k^{(1)}$ in (I.0.5), it holds $\forall k \in \mathcal{K}$,

$$\mathbb{E}\left[V_k^{(1)}(n)\,\middle|\,\mathcal{F}^{(1)}_{k,n-1}\right] \overset{n\to\infty}{\longrightarrow} 0. \tag{I.0.19}$$

□

From *Lemmas I.0.4, I.0.5, I.0.6* and *Theorem* 2 *in [14]*, we have that the inhomogeneous, discrete, random time process $\{\hat{\mathbf{u}}_k(n)\}_{n\in\mathbb{N}}$ in (I.0.6) asymptotically approaches the autonomous, continuous, deterministic time dynamical system $\widehat{\mathbf{u}}_k(t)$, $t \in [0,\infty[$ determined by the ordinary differential equation (ODE),

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{u}}_k(t) = \mathbf{U}_k^{(1)}\left(\widehat{\mathbf{u}}_k(t)\right), t \in [0,\infty[, \tag{I.0.20}$$

and the initialization values $\widehat{\mathbf{u}}_k(0)$, $\forall k \in \mathcal{K}$. In the following, we arbitrarily assume that $\widehat{\mathbf{u}}_k(0) = \hat{\mathbf{u}}_k(0)$. Now, focus on the n_k-th component of the ODE (I.0.20) of player k, with $k \in \mathcal{K}$, and $n_k \in \{1,\ldots,N_k\}$, i.e.,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{u}}_{k,n_k}(t) &= U_{k,n_k}^{(1)}\left(\widehat{\mathbf{u}}_k(t)\right), \\ &= \bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*\right) - \widehat{u}_{k,A_k^{(n_k)}}(t).\end{aligned} \tag{I.0.21}$$

The solution of the ODE (I.0.21) is,

$$\widehat{u}_{k,n_k}(t) = \bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*\right) - \left(\bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*\right) - \hat{u}_{k,A_k^{(n_k)}}(0)\right) e^{-\left(\pi^+_{k,A_k^{(n_k)}}\right)t}, \; t \in [0,\infty[. \tag{I.0.22}$$

From (I.0.22), it follows that,

$$\lim_{t\to\infty} \widehat{u}_{k,n_k}(t) = \bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^*\right), \tag{I.0.23}$$

which proves that if $\hat{u}_{k,A_k^{(n_k)}}$ follows the time-averaging rule (4.29), Lemma 4.5.1 holds.

Now, we focus, on the iterative algorithm of the form

$$\hat{u}_{k,A_k^{(n_k)}}(n) = \hat{u}_{k,A_k^{(n_k)}}(n-1) + \alpha_k(n)\frac{\mathbb{1}_{\left\{a_k(n)=A_k^{(n_k)}\right\}}}{\pi^+_{k,A_k^{(n_k)}}}\left(\tilde{u}_{k(n)} - \hat{u}_{k,A_k^{(n_k)}(n-1)}\right) \tag{I.0.24}$$

Let the function $G_k^{(2)} : \mathcal{A}_k \times \mathbb{R}^{N_k+1} \to \mathbb{R}^{N_k}$ be defined as follows for all $k \in \mathcal{K}$,

$$G_k^{(2)}(a_k,\hat{\mathbf{u}}_k,\tilde{u}_k) = \left(\frac{\mathbb{1}_{\left\{a_k=A_k^{(1)}\right\}}}{\pi^+_{k,A_k^{(1)}}}\left(\tilde{u}_k - \hat{u}_{k,A_k^{(1)}}\right),\ldots,\frac{\mathbb{1}_{\left\{a_k=A_k^{(1)}\right\}}}{\pi^+_{k,A_k^{(N_k)}}}\left(\tilde{u}_k - \hat{u}_{k,A_k^{(N_k)}}\right)\right), \tag{I.0.25}$$

where, $a_k \in \mathcal{A}_k$, $\hat{\mathbf{u}}_k \in \mathbb{R}^{N_k}$, and $\tilde{u}_k \in \mathbb{R}$, for all $n > 0$. Let also the function $U_k^{(2)} : \mathbb{R}^{N_k} \to \mathbb{R}^{N_k}$ be defined as follows,

$$\mathbf{U}_k^{(2)}(\hat{\mathbf{u}}_k(n)) = \mathbb{E}_{\mathbf{h},\pi^+,\varepsilon_k}\left[G^{(2)}(a_k, \hat{\mathbf{u}}_k(n), u_k(\mathbf{h}, a_k, \mathbf{a}_{-k}) + \varepsilon_{k,a_k})\right]. \tag{I.0.26}$$

Denote by $U_{k,n_k}^{(2)}(\hat{\mathbf{u}}_k(n))$, with $n_k \in \{1,\ldots,N_k\}$, the n_k-th component of the vector $\mathbf{U}_k^{(2)}(\hat{\mathbf{u}}_k(n))$ such that, $\mathbf{U}_k^{(2)}(\hat{\mathbf{u}}_k) = \left(U_{k,1}^{(1)}(\hat{\mathbf{u}}_k),\ldots,U_{k,N_k}^{(1)}(\hat{\mathbf{u}}_k)\right)$ and for all $n_k \in$

$\{1,\ldots,N_k\}$, it follows from (I.0.26) that,

$$\begin{aligned}
U_{k,n_k}^{(2)}(\hat{\mathbf{u}}_k) &= \mathbb{E}_{\mathbf{h},\pi^*,\varepsilon_k}\left[\frac{\mathbb{1}_{\left\{a_k=A_k^{(n_k)}\right\}}}{\pi^+_{k,A_k^{(n_k)}}}\left(u_k\left(\mathbf{h},A_k^{(n_k)},\mathbf{a}_{-k}\right)+\varepsilon_{k,A_k^{(n_k)}}-\hat{u}_{k,A_k^{(n_k)}}\right)\right]\\
&= \mathbb{E}_{\mathbf{h},\pi^*,\varepsilon_k}\left[\frac{\mathbb{1}_{\left\{a_k=A_k^{(n_k)}\right\}}}{\pi^+_{k,A_k^{(n_k)}}}\left(u_k\left(\mathbf{h},A_k^{(n_k)},\mathbf{a}_{-k}\right)+\varepsilon_{k,A_k^{(n_k)}}-\hat{u}_{k,A_k^{(n_k)}}\right)\right]\\
&= \frac{1}{\pi^+_{k,A_k^{(n_k)}}}\mathbb{E}_{\pi_k^*}\left[\mathbb{1}_{\left\{a_k=A_k^{(n_k)}\right\}}\right]\left(\mathbb{E}_{\mathbf{h},\pi^*_{-k}}\left[u_k\left(\mathbf{h},A_k^{(n_k)},\mathbf{a}_{-k}\right)\right]+\right.\\
&\qquad \left.\mathbb{E}_{\varepsilon_k}\left[\varepsilon_{k,A_k^{(n_k)}}\right]-\hat{u}_{k,A_k^{(n_k)}}\right)\\
&= \bar{u}_k\left(\mathbf{e}_{n_k}^{(N_k)},\pi^*_{-k}\right)-\hat{u}_{k,A_k^{(n_k)}}. \qquad\qquad (I.0.27)
\end{aligned}$$

Define the function $V_k^{(2)}:\mathbb{N}\setminus\{0\}\to\mathbb{R}$ as follows,

$$\mathbf{V}_k^{(2)}(n) \;=\; G_k^{(2)}(a_k(n),\hat{\mathbf{u}}_k(n-1),\tilde{u}_k(n))-U_k^{(2)}(\hat{\mathbf{u}}_k(n-1)). \qquad (I.0.28)$$

Hence, using the equations (I.0.26) and (I.0.28), we can write equation (I.0.1) as follows.

$$\hat{\mathbf{u}}_k(n)=\hat{\mathbf{u}}_k(n-1)+\alpha_k(n)\left(U_k^{(2)}\left(\hat{\mathbf{u}}_k(n-1)\right)-V_k^{(2)}(n)\right). \qquad (I.0.29)$$

Now, we introduce the following lemmas.

Lemma I.0.7 *For all $k\in\mathcal{K}$, the function $U_k^{(2)}$ is Lipschitz with Lipschitz constant $L=1$.*

The proof of Lemma I.0.7 follows the same steps of the proof of Lemma I.0.7.

Lemma I.0.8 *The sequence $\{\mathbf{V}^{(2)}(n)_k\}_{n>0}$ is a martingale difference sequence with respect to the increasing family of σ-fields,*

$$\mathcal{F}_{k,n}^{(2)}\triangleq\sigma\left(\left(\hat{\mathbf{u}}_k(0),\pi(0),V_k^{(2)}(0),\mathbf{h}(0),\varepsilon(0)\right),\ldots,\left(\hat{\mathbf{u}}_k(n),\pi(n),V_k^{(2)}(n),\mathbf{h}(n),\varepsilon(n)\right)\right), \qquad (I.0.30)$$

The proof of Lemma I.0.8 follows the same steps of the proof of Lemma I.0.6.
From *Lemmas I.0.7, I.0.8* and the conditions in (4.31), and *Theorem* 2 *in [14]*, we have that the inhomogeneous, discrete, random time process $\{\hat{\mathbf{u}}_k(n)\}_{n\in\mathbb{N}}$ in (I.0.24) asymptotically approaches the autonomous, continuous, deterministic time dynamical system $\hat{\mathbf{u}}_k(t)$, $t\in[0,\infty[$ determined by the ordinary differential equation (ODE),

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{u}}_k(t)=\mathbf{U}_k^{(2)}\left(\hat{\mathbf{u}}_k(t)\right),t\in[0,\infty[, \qquad (I.0.31)$$

and the initialization values $\hat{\mathbf{u}}_k(0)$, $\forall k\in\mathcal{K}$. In the following, we arbitrarily assume that $\hat{\mathbf{u}}_k(0)=\hat{\mathbf{u}}_k(0)$. Now, focus on the n_k-th component of the ODE (I.0.31) of player k, with $k \in \mathcal{K}$ and $n_k \in \{1, \ldots, N_k\}$, i.e.,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\mathbf{u}}_{k,n_k}(t) &= U^{(1)}_{k,n_k}\left(\widehat{\mathbf{u}}_k(t)\right), \\ &= \bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^*_{-k}\right) - \widehat{u}_{k,A_k^{(n_k)}}(t). \end{aligned} \quad \text{(I.0.32)}$$

The solution of the ODE (I.0.32) is,

$$\widehat{u}_{k,n_k}(t) = \bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^*_{-k}\right) - \left(\bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^*_{-k}\right) - \hat{u}_{k,A_k^{(n_k)}}(0)\right)e^{-t},\ t \in [0, \infty[. \quad \text{(I.0.33)}$$

From (I.0.33), it follows that,

$$\lim_{t\to\infty} \widehat{u}_{k,n_k}(t) = \bar{u}_k\left(\mathbf{e}^{(N_k)}_{n_k}, \pi^*_{-k}\right). \quad \text{(I.0.34)}$$

Finally, since $\hat{\mathbf{u}}_k(n)$ asymptotically tracks $\widehat{u}_{k,n_k}(t)$, we have proved Lemma 4.5.1

Appendix J

Proof of Theorem 4.5.2

In this appendix, we present the proof of Theorem 4.5.2. As a first step, we recall that in Lemma 4.5.1 it has been shown that

$$\lim_{t\to\infty} ||\hat{\mathbf{u}}_k(\lfloor t \rfloor) - \hat{\mathbf{u}}_k(t)|| = 0, \tag{J.0.1}$$

where $\forall x \in \mathbb{R}_+$, such that $n \leqslant x \leqslant n+1$, with $n \in \mathbb{N}$, $\lfloor x \rfloor = n$. Thus, for a sufficiently large $t > 0$, the following equality can be accepted,

$$\eta_{k,A_k^{(n_k)}}(t) \;=\; \left|\widehat{u}_{k,A_k^{(n_k)}}(t) - \bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^{+})\right| \tag{J.0.2}$$

Then, from the solution of the ODE (I.0.20) given by (I.0.22), we have that

$$\eta_{k,A_k^{(n_k)}}(t) \;=\; \left|\bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^{+}) - \hat{u}_{k,A_k^{(n_k)}}(0)\right| \exp\left(-\left(\pi^{+}_{k,A_k^{(n_k)}}\right)t\right). \tag{J.0.3}$$

Now from the condition $\eta_{k,A_k^{(n_k)}}(t) < \eta$, it yields,

$$t > \frac{1}{\pi^{+}_{k,A_k^{(n_k)}}} \ln\left(\frac{\left|\bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^{+}) - \hat{u}_{k,A_k^{(n_k)}}(0)\right|}{\eta}\right). \tag{J.0.4}$$

Similarly, from the solution to the ODE (I.0.31) given by (I.0.33), we have that

$$\eta_{k,A_k^{(n_k)}}(t) \;=\; \left|\bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^{+}) - \hat{u}_{k,A_k^{(n_k)}}(0)\right| \exp\left(-t\right). \tag{J.0.5}$$

From the condition $\eta_{k,A_k^{(n_k)}}(t) < \eta$, it yields,,

$$t > \ln\left(\frac{\left|\bar{u}_k(\mathbf{e}_{n_k}^{(N_k)}, \pi_{-k}^{+}) - \hat{u}_{k,A_k^{(n_k)}}(0)\right|}{\eta}\right), \tag{J.0.6}$$

which ends the proof of Theorem 4.5.2.

Appendix K

Proof of Lemma 4.5.3

In this appendix, we provide the proof of Lemma 4.5.3. This proof is divided in two parts. First, we show that the learning dynamics in (4.38) can be approximated by a system of ODEs. Second, we show that that the zeros of such system of ODEs corresponds to a Logit Equilibrium.

Part i**: *ODE Approximation*.**
The learning dynamics in (4.38) under assumption ($B1$) can be written, for all $k \in \mathcal{K}$, as follows,

$$\pi_k(n) \quad = \quad \pi_k(n-1) + \lambda_k(n) \left(\beta_k \left(\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n)) \right) - \pi_k(n-1) \right). \qquad \text{(K.0.1)}$$

Let the function $\mathbf{U}_k^{(3)} : \mathbb{R}^{\sum_{j=1}^{K} N_j} \rightarrow \mathbb{R}^{N_k}$, with $k \in \mathcal{K}$, be defined as follows,

$$\mathbf{U}_k^{(3)} \left(\pi_k, \pi_{-k} \right) = \beta_k \left(\bar{\mathbf{u}}_k(\cdot, \pi_{-k}(n)) \right) - \pi_k(n-1). \qquad \text{(K.0.2)}$$

Note that for all $n > 0$, it holds that $||\pi_k(n)|| < \infty$, since it is a probability distribution, and $||u_k(n)|| < \infty$ since it is the mean of finite values $\tilde{u}_k(0), \ldots, \tilde{u}_k(n-1)$. Note also that the functions β_k and $\bar{u}_k$ are continous over the spaces $\mathbb{R}^{N_k}$ and $\triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$, respectively. Thus, the function $\mathbf{U}_k^{(3)}$ is continous with bounded gradient at least over the space $\triangle(\mathcal{A}_1) \times \ldots \times \triangle(\mathcal{A}_K)$. Hence, the function $\mathbf{U}_k^{(3)}$ is Lipchitz. Now, since for all $k \in \mathcal{K}$ and $n > 0$, the learning rates follows the condition in (4.37), it holds from *Theorem* 2 *in [14]* that the coupled learning dynamics in (4.38) can be approximated by the system of ODEs,

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_1(t) & = & \beta_1 \left(\bar{\mathbf{u}}_1 \left(\cdot, \widehat{\pi}_{-1}(t) \right) \right) - \widehat{\pi}_1(t), \\ \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_2(t) & = & \beta_2 \left(\bar{\mathbf{u}}_2 \left(\cdot, \widehat{\pi}_{-2}(t) \right) \right) - \widehat{\pi}_2(t), \\ & \vdots & \\ \frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_K(t) & = & \beta_K \left(\bar{\mathbf{u}}_K \left(\cdot, \widehat{\pi}_{-K}(t) \right) \right) - \widehat{\pi}_K(t). \end{cases} \qquad \text{(K.0.3)}$$

Part ii**: *The Zeros of the system of ODEs*.**
Note that if convergence occurs, then the limit point π^* must be a zero of the system of ODEs (K.0.3). Hence, it follows that $\forall k \in \mathcal{K}$,

$$\frac{\mathrm{d}}{\mathrm{d}t}\widehat{\pi}_k(t) = 0, \qquad \text{(K.0.4)}$$

implies that

$$\pi_k^* = \beta_k \left(\bar{\mathbf{u}}_k \left(\cdot, \pi_{-k}^* \right) \right), \tag{K.0.5}$$

which corresponds to the definition of LE (Def. 4.4.2). This completes the proof of Lemma 4.5.3.